



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number 190532

**TO: Rosanne Kosson**  
**Art Unit: 1653**  
**Location: REM-3B84&3C18**  
**Serial Number: 09/856723**

**Tuesday, May 23, 2006**

**From: Beverly Shears**  
**Location: Biotech-Chem Library**  
**REM 1A54**  
**Phone: 571-272-2528**  
**beverly.shears@uspto.gov**

### Search Notes

Your queries have completed processing. You may access an electronic version via eDAN (SCORE) and /or <http://es/ScoreAccessWeb>. If the result files have been separated into two (2) or more versions, you may view additional files via the select "[View version list for this application](#)" link.

#### Protein Sequence Searches – February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (uniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

#### Published Applications Database - November 2005

Published\_Applications Nucleic Acid and Published\_Applications Amino Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).



**THIS PAGE BLANK (USPTO)**

STIC Biotech/ChemLib

190532

From: Kosson, Rosanne  
Sent: Thursday, May 18, 2006 3:55 PM  
To: STIC-Biotech/ChemLib  
Subject: request for sequence search- 09/856,723- REGULATORY PROTEIN PKE#83 FROM HUMAN KERATINOCYTES

Please search SEQ ID NO: 8 for me in the commercial and interference databases. This is a protein of 1076 amino acids. My mailbox is in 3C70. Thanks!!!

Rosanne Kosson  
Patent Examiner, AU 1653  
REM 3B84  
571-272-2923  
rosanne.kosson@uspto.gov

803002

CRFE

aa 1076

PD = 11/19/99 PCT  
12/7/98 DE ? no  
11/26/98 DE ? no

pct/de 99/03732

RECEIVED  
MAY 18 2006  
STIC-BIOTECH/CHM LIB

ME

5/12  
IT

\*\*\*\*\*

Searcher: \_\_\_\_\_  
Searcher Phone: \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date completed: \_\_\_\_\_  
Searcher Prep Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search  
NA# \_\_\_\_\_ AA# \_\_\_\_\_  
S/L: \_\_\_\_\_ Oligomer: \_\_\_\_\_  
Encode/Transl: \_\_\_\_\_  
Structure #: \_\_\_\_\_ Text: \_\_\_\_\_  
Inventor: \_\_\_\_\_ Litigation: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (Specify): \_\_\_\_\_

Date completed: \_\_\_\_\_

Searcher: Beverly e 2528

Terminal time: \_\_\_\_\_

Elapsed time: \_\_\_\_\_

CPU time: \_\_\_\_\_

Total time: \_\_\_\_\_

Number of Searches: \_\_\_\_\_

Number of Databases: \_\_\_\_\_

Search Site

\_\_\_\_\_ STIC

\_\_\_\_\_ CM-1

\_\_\_\_\_ Pre-S

Type of Search

\_\_\_\_\_ N.A. Sequence

\_\_\_\_\_ A.A. Sequence

\_\_\_\_\_ Structure

\_\_\_\_\_ Bibliographic

Vendors

\_\_\_\_\_ IG

\_\_\_\_\_ STN

\_\_\_\_\_ Dialog

\_\_\_\_\_ APS

\_\_\_\_\_ Geninfo

\_\_\_\_\_ SDC

\_\_\_\_\_ DARC/Questel

\_\_\_\_\_ Other CGN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using BW model

Run on: May 20, 2006, 17:05:42 ; Search time 320 Seconds

(without alignments)  
3110.365 Million cell updates/sec

Title: US-09-856-723A-8

Sequence: 1 MKQYASPMPTQTDVTKFKP.....RTLBNKGMKAKKEKCVLQ 1076

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : UniProt 7.2.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1196	2	Q6QNV2_HUMAN
2	5442.5	99.2	1231	2	Q6ND11_HUMAN
3	5265	95.9	1160	2	Q6N372_HUMAN
4	4919	89.6	962	2	Q69J77_HUMAN
5	4683.5	85.3	1242	2	Q692W3_MOUSE
6	4552	82.9	1206	2	Q6SOK3_MOUSE
7	4532	82.6	1206	2	Q91ZJ6_MOUSE
8	4414.5	80.4	893	2	Q51TV6_HUMAN
9	2255.5	41.3	1036	2	Q4SGB9_TRTNG
10	1154.5	21.0	762	2	Q5PQW3_FAT
11	1151.5	21.0	874	2	Q99J97_MOUSE
12	1131	20.6	759	2	Q99MS6_MOUSE
13	1127	20.5	771	2	Q99MS5_MOUSE
14	1082	19.7	723	2	Q8C2W6_MOUSE
15	1076.5	19.6	987	2	Q9V7X1_DROME
16	1075.5	19.6	987	2	Q6MR33_DROME
17	1018.5	18.6	987	2	Q7PKB6_ANOGA
18	966	17.6	310	2	Q9MW19_HUMAN
19	950.5	17.3	1716	2	Q99MS7_MOUSE
20	945.5	17.2	1415	2	Q9H7M7_HUMAN
21	941.5	17.2	1510	2	Q8N3D4_HUMAN
22	895	16.3	308	2	Q5SQK2_MOUSE
23	850.5	15.5	1919	2	Q4R1P0_TRTNG
24	627	11.4	125	2	Q53T67_HUMAN
25	579.5	10.6	288	2	Q4SGC0_TRTNG
26	556	10.1	243	2	Q8TB89_HUMAN
27	507.5	9.2	1010	2	Q9VU34_DROME
28	498.5	9.1	863	1	MILK1_HUMAN
29	480.5	8.8	967	1	Q6GN09_XENLA
30	480	8.7	870	1	MILK1_MOUSE
31	467.5	8.5	1886	2	Q47H5_TRTNG

32	453.5	8.3	1190	2	Q519A0_ENTH1
33	452	8.2	211	2	Q580X2_HUMAN
34	434	7.9	945	2	Q60W15_CABER
35	431.5	7.9	904	1	MILK2_HUMAN
36	418.5	7.6	1009	2	Q3UB65_MOUSE
37	416.5	7.6	1009	2	Q3UN34_MOUSE
38	414.5	7.3	1009	2	Q3TPY3_MOUSE
39	402.5	7.3	901	2	Q18749_CABER
40	402	7.3	583	2	Q4SMU9_TRTNG
41	389.5	7.1	992	2	Q5DU62_MOUSE
42	379.5	6.9	242	2	Q4T0T5_TRTNG
43	368.5	6.7	989	2	Q94909_HUMAN
44	354.5	6.5	1738	2	Q76329_DICDI
45	344	6.3	2206	2	Q4REQ4_TRTNG

## ALIGNMENTS

Query Match	Score	Length	ID	Description
Best Local Similarity	99.9%	Score 5486; DB 2; Length 1196;		
Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;				
QY	1	MKQYASPMPTQTDVTKFKPDKSKVSAALQFSISCTFLBKGKATDEDMOSLASVSMQ 60	Q6QNV2_HUMAN	PRELIMINARY; PRT; 1196 AA.
DB	121	MKQYASPMPTQTDVTKFKPDKSKVSAALQFSISCTFLBKGKATDEDMOSLASVSMQ 180	Q6QNV2; Q6QNV2;	
QY	61	ADIGNLDPFEDNDDEENRVOEERAKITTELINKNLPFDEAKDIAVNSNFPDDPA 120	Q6QNV2; Q6QNV2;	
DB	181	ADIGNLDPFEDNDDEENRVOEERAKITTELINKNLPFDEAKDIAVNSNFPDDPA 240	Q6QNV2; Q6QNV2;	
QY	121	AEINLPDPSSEETITETASPRKTEDSFYNSYNPFKEVOTPOYLNFDEBEAFVTKDS 180	Q6QNV2; Q6QNV2;	
DB	241	AEINLPDPSSEETITETASPRKTEDSFYNSYNPFKEVOTPOYLNFDEBEAFVTKDS 300	Q6QNV2; Q6QNV2;	
QY	181	PPGSTKKNIRPVDMSKYLADSSKTEEBELDSNPYBKSIPPNNLVNPOVLETER 240	Q6QNV2; Q6QNV2;	
DB	301	PPGSTKKNIRPVDMSKYLADSSKTEEBELDSNPYBKSIPPNNLVNPOVLETER 360	Q6QNV2; Q6QNV2;	
QY	241	RVRKKAAPAVLSPKTCVNLNENTVSAQKDLSTSPKPSPIPSVLGRKPNASQSLVWCKE 300	Q6QNV2; Q6QNV2;	

```

Db 361 RYKAKAPAPVLSKPTGLNENTVSAGKDLSTSPKSPSPAPVLRGRKPNASQSLVWCKE 420
Qy 301 VTKNRYGVKLTNTFTTSWRNGLSFCAIILHFRPDLIDYSLNPDIKENNKAYGFGASIG 360
Db 421 VTKNRYGVKLTNTFTTSWRNGLSFCAIILHFRPDLIDYSLNPDIKENNKAYGFGASIG 480
Qy 361 ISRLLESMDVLLAIIPDKLTMTYLYQIRAHFSGOELNVVQIIEENSSKSTKYGVNETD 420
Db 481 ISRLLESMDVLLAIIPDKLTMTYLYQIRAHFSGOELNVVQIIEENSSKSTKYGVNETD 540
Qy 421 NSSVDQEKYFAELSDLKREPELQOPISGAVDFLSQDSDVPVNDGSGVGESESHQTPDHL 480
Db 541 NSSVDQEKYFAELSDLKREPELQOPISGAVDFLSQDSDVPVNDGSGVGESESHQTPDHL 600
Qy 481 SPSTASPYCRRTKSDTEPQKSGQSSGRTSGSDPGICSNMTDSTQAQVLLGKRLKAEYL 540
Db 601 SPSTASPYCRRTKSDTEPQKSGQSSGRTSGSDPGICSNMTDSTQAQVLLGKRLKAEYL 660
Qy 541 ELSDLYSDKKDKMSPPICETDEOKLQTLIDISNLEKEKLENSRSLGCRSDPESPIK 600
Db 661 ELSDLYSDKKDKMSPPICETDEOKLQTLIDISNLEKEKLENSRSLGCRSDPESPIK 720
Qy 601 TSLSPTSKLGYSYSRDLIDLAKKKKASLRQTESDPDADRTTLNHADHSSKIYQHRLLSRQ 660
Db 721 TSLSPTSKLGYSYSRDLIDLAKKKKASLRQTESDPDADRTTLNHADHSSKIYQHRLLSRQ 780
Qy 661 ELKERAVLLEQARDAALAKGNHNTATPCRCNRLSDQDERRRQLERAROLIAE 720
Db 781 ELKERAVLLEQARDAALAKGNHNTATPCRCNRLSDQDERRRQLERAROLIAE 840
Qy 721 ARSGVKMSELPSYGEAAEKLKERSKASGDENDNIETDNEEIPGFGVGGDELTYLEN 780
Db 841 ARSGVKMSELPSYGEAAEKLKERSKASGDENDNIETDNEEIPGFGVGGDELTYLEN 900
Qy 781 DLDTPEQNSKLVLDKLEKLEVPQVANSPPSAQAQVTESSBQDMKSGTEDLATERLQK 840
Db 901 DLDTPEQNSKLVLDKLEKLEVPQVANSPPSAQAQVTESSBQDMKSGTEDLATERLQK 960
Qy 841 TTRFRFPVYVSKOSTYRKTQLOFSQYIENRPMKQORSIQETTKGNEKKAITYTOR 900
Db 961 TTRFRFPVYVSKOSTYRKTQLOFSQYIENRPMKQORSIQETTKGNEKKAITYTOR 1020
Qy 901 KPSEDEVLANGFQKOTSGYVVGELALLENQOKIDTRAAVLEKRLRYLMDGRNTESEAM 960
Db 1021 KPSEDEVLANGFQKOTSGYVVGELALLENQOKIDTRAAVLEKRLRYLMDGRNTESEAM 1080
Qy 961 MOEWFMVLNKKNALIRRMNQSLLEKEHDLERRYELNRELRAVLAIEDMQTEAQKRE 1020
Db 1081 MOEWFMVLNKKNALIRRMNQSLLEKEHDLERRYELNRELRAVLAIEDMQTEAQKRE 1140
Qy 1021 QLLDELVALVNRGDALVRDLDAQEKQAEDEDEHLETTLEONKGMKAKKEKCYLQ 1076
Db 1141 QLLDELVALVNRGDALVRDLDAQEKQAEDEDEHLETTLEONKGMKAKKEKCYLQ 1196

```

## RESULT 2

```

O8ND11_HUMAN PRELIMINARY; PRT; 1231 AA.
AC O8ND11;
DT 01-OCT-2002, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 17.
DE Hypothetical protein DKFZ566N1646 (Npr/calponin-like protein).
GN Name=EHBP1; Synonym=DKFZ566N1646, NACSIN;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Kidney;

```

```

RG The German cDNA Consortium;
RA Oettermajer B., Oettermajer B., Deutschenbaur S., Schajp A.,
RA Mewes H.W., Weil B., Amid C., Oeangner A., Fodor G., Han M., Wiemann S.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RA Castellano-Munoz M., Fernandez-Chacon R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL: AL833968; CAD38814.1; -; mRNA.
DR EMBL: AY311386; AAQ97141.1; -; mRNA.
DR HSP: O01082; 1BKX.
DR Ensembl: ENSG00000115504; Homo sapiens.
DR HGNC: HGNC:29144; EHBP1.
DR InterPro: IPR001715; Calponin_act_bd.
DR Pfam: PF00307; CH; 1.
DR SMART: SM00033; CH; 1.
DR PROSITE: PSS0021; CH; 1.
KW Hypothetical protein.
SQ SEQUENCE 1231 AA; 139658 MW; DB7A7DBE672AEB98 CRC64;

Query Match 99.2%; Score 5442.5; DB 2; Length 1231;
Best Local Similarity 96.6%; Pred. No. 4,1e-205;
Matches 1073; Conservative 1; Mismatches 2; Indels 35; Gaps 1;

Qy 1 MKQYASPMPTQTDVKKFKPLSKKVVSAALQFSLSCTFLREGKATBEDMQSLASVSMQ 60
Db 121 MKQYASPMPTQTDVKKFKPLSKKVVSAALQFSLSCTFLREGKATBEDMQSLASVSMQ 180
Qy 61 ADIGNLDDPEEDNEDDENRVNQEERAKIT----- 91
Db 181 ADIGNLDDPEEDNEDDENRVNQEERAKITGIVNQLALSLDEDDQDCIKANRSK 240
Qy 92 -----ELINKLNFLEAEKDLATVNSNPFDDPAELNPFQDPDSEPTETASPRKTE 145
Db 241 SASSSEELINKLNFLEAEKDLATVNSNPFDDPAELNPFQDPDSEPTETASPRKTE 300
Qy 146 DSFYNNSYNPFKEVQTPQYLNPPDEEAFVTIKDSPQSTKRNIRPVMKXYLVADSK 205
Db 301 DSFYNNSYNPFKEVQTPQYLNPPDEEAFVTIKDSPQSTKRNIRPVMKXYLVADSK 360
Qy 206 TEREELDESNPFPKSTPPNNLVNVPQELTEREVKKAAPVLSKPTGLNENTVS 245
Db 361 TEREELDESNPFPKSTPPNNLVNVPQELTEREVKKAAPVLSKPTGLNENTVS 420
Qy 245 TEREELDESNPFPKSTPPNNLVNVPQELTEREVKKAAPVLSKPTGLNENTVS 300
Db 420 TEREELDESNPFPKSTPPNNLVNVPQELTEREVKKAAPVLSKPTGLNENTVS 480
Qy 325 AGKCLSTSPKSPIPSPVLRGRKPNASQSLVWCKEYTKNRYGVKLTNTFTTSWRNGLSFCA 385
Db 480 AGKCLSTSPKSPIPSPVLRGRKPNASQSLVWCKEYTKNRYGVKLTNTFTTSWRNGLSFCA 540
Qy 385 ILHFRPDLIDYKSLNPDIKENNKAYGFGASIGISRLLESMDVLLAIIPDKLTMTYLY 445
Db 540 ILHFRPDLIDYKSLNPDIKENNKAYGFGASIGISRLLESMDVLLAIIPDKLTMTYLY 600
Qy 445 ISGAVDFLSQDSDVPVNDGSGVGESESHQTPDHLSPSTASPYCRRTKSDTEPQKSGQSS 505
Db 600 ISGAVDFLSQDSDVPVNDGSGVGESESHQTPDHLSPSTASPYCRRTKSDTEPQKSGQSS 660
Qy 505 GRTSGSDPGICSNMTDSTQAQVLLGKRLKAEYLSDLYVSDKKDKMSPPICETDE 565
Db 660 GRTSGSDPGICSNMTDSTQAQVLLGKRLKAEYLSDLYVSDKKDKMSPPICETDE 720
Qy 565 OKLQTLIDISNLEKEKLENSRSLGCRSDPESPIKTSLSPTSKLGYSYSRDLIDLAKKKA 625
Db 720 OKLQTLIDISNLEKEKLENSRSLGCRSDPESPIKTSLSPTSKLGYSYSRDLIDLAKKKA 780
Qy 625 SLRQTESDPDADRTTLNHADHSSKIYQHRLLSRQELKERAVLLEQARDAALAKGNHG 685

```

Db 781 SLRQTESPDDADRTLTNADHSSKIVQRLRLSROBELERARVLLLEQARRDALAKGNKH 840  
Qy 686 NTNTATPCNRLQSLQOQDEERRRQURERARQLIABASGVMSLPSYGEAAETLKRSS 745  
Db 841 NTNTATPCNRLQSLQOQDEERRRQURERARQLIABASGVMSLPSYGEAAETLKRSS 900  
Qy 746 KASGDENDNIDITNEEIPSGFVVGGDELNTLENDLTPPONSGLVYLKLLLEVOPO 805  
Db 901 KASGDENDNIDITNEEIPSGFVVGGDELNTLENDLTPPONSGLVYLKLLLEVOPO 960  
Qy 806 VANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQKTERFRNPVVSFSDSTVRKQLQGSF 865  
Db 961 VANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQKTERFRNPVVSFSDSTVRKQLQGSF 1020  
Qy 866 SOYIENPREMKRQSIQDITKKGNEKKAITETQKSEEDSVLANKGFOTQYVVGELAA 925  
Db 1021 SOYIENPREMKRQSIQDITKKGNEKKAITETQKSEEDSVLANKGFOTQYVVGELAA 1080  
Qy 926 LENEQOQIDTAAALVEKRLRYLMDTGRNTEEBEAMQEMFLVNGKALIRRNQSLLE 985  
Db 1081 LENEQOQIDTAAALVEKRLRYLMDTGRNTEEBEAMQEMFLVNGKALIRRNQSLLE 1140  
Qy 986 KEHDLERRYELLNRELRLAMLAIEDWQKTEAQKREQLLDELVALVNRDALVVDLDAQE 1045  
Db 1141 KEHDLERRYELLNRELRLAMLAIEDWQKTEAQKREQLLDELVALVNRDALVVDLDAQE 1200  
Qy 1046 KOABEBDEHLERTLEQNGKMAKKEKCVLQ 1076  
Db 1201 KOABEBDEHLERTLEQNGKMAKKEKCVLQ 1231

RESULT 3  
06NKT2\_HUMAN  
ID 06NKT2\_HUMAN PRELIMINARY; PRT; 1160 AA.  
AC 06NKT2;  
DT 05-JUL-2004, integrated into UniProtKB/TrEMBL.  
DT 05-JUL-2004, sequence version 1.  
DT 07-FEB-2006, entry version 13.  
DE EMBL protein.  
GN Name=EBP1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;  
OC Homo.  
NCBI TaxID=9606;  
RN NCBI TaxID=9606;  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Uterus;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stieplon M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein W.J., Usdin T.B., Toshiyuki S., Carrinci F., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton B., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smallus D.E.,  
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences."  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC TISSUE=Uterus;

RA Director MGC Project;  
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
CC Distributed under the Creative Commons Attribution-NonDerivs License  
CC -----  
DR EMBL; BC067215; AAH67215.1; -; mRNA.  
DR HSSP; Q01082; 1AA2.  
DR InterPro; IPR001715; Calponin\_act\_bd.  
DR Pfam; PF00307; CH; 1.  
DR SMART; SM00033; CH; 1.  
DR PROSITE; PS50021; CH; 1.  
SQ SEQUENCE 1160 AA; 132322 MW; DD79233FE0B14AF1 CRC64;  
Query Match 95.9%; Score 5266; DB 2; Length 1160;  
Beet Local Similarity 96.6%; Pred. No. 3,2e-198;  
Matches 1039; Conservative 1; Mismatches 0; Indels 36; Gaps 1;  
Qy 1 MKOYASPMPTQTDVAKLKFPLSKKYVSAALQFSLSCIFLREGKATDEDMQSLASVSKQ 60  
Db 121 MKOYASPMPTQTDVAKLKFPLSKKYVSAALQFSLSCIFLREGKATDEDMQSLASVSKQ 180  
Qy 61 ADIGNLDDFEEDNEDDENRVNQEKAKITELINKLFLDEAKDLATVNSNPPDDPA 120  
Db 181 ADIGNLDDFEEDNEDDENRVNQEKAKITELINKLFLDEAKDLATVNSNPPDDPA 240  
Qy 121 AELNPPDDPSEEEITETASPRKTEDSFYNSNYPFEVQTPQYLPNPPDEAFVTIKDS 180  
Db 241 AELNPPDDPSEEEITETASPRKTEDSFYNSNYPFEVQTPQYLPNPPDEAFVTIKDS 300  
Qy 181 PPGSTKRNIRPVMSKYLVADSSKTEEBELDESNPFEKSTPPNNLVVPOVLETER 240  
Db 301 PPGSTKRNIRPVMSKYLVADSSKTEEBELDESNPFEKSTPPNNLVVPOVLETER 360  
Qy 241 RVKKKAPAPVLSPTKVLNMENTVYSAKDLSTSPKSPSPSPVLRKRNASQSLVWCKE 300  
Db 361 RVKKKAPAPVLSPTKVLNMENTVYSAKDLSTSPKSPSPSPVLRKRNASQSLVWCKE 420  
Qy 361 VTKNYRGVKTNTFTTSNRNGLSFCAIIHHRPDLIDYKSLNPQIKENKKAAYGFSIG 360  
Db 421 VTKNYRGVKTNTFTTSNRNGLSFCAIIHHRPDLIDYKSLNPQIKENKKAAYGFSIG 480  
Qy 361 ISRLBESDMVLAIAPKLITMTLYQIRAHFSGQELNVQIBNSSKSTYKVGNYETD 420  
Db 481 ISRLBESDMVLAIAPKLITMTLYQIRAHFSGQELNVQIBNSSKSTYKVGNYETD 540  
Qy 481 NSSVDOQKFTYAEISDLKREBELQPIGAVDFLSQDDSVFVNSGVESESEHQTPDHL 480  
Db 541 NSSVDOQKFTYAEISDLKREBELQPIGAVDFLSQDDSVFVNSGVESESEHQTPDHL 600  
Qy 481 SPSTASPYCRRTYSDETPQKSSQSSGRTSGSDPGICSNSTDSQAQVLLGKRLKATL 540  
Db 601 SPSTASPYCRRTYSDETPQKSSQSSGRTSGSDPGICSNSTDSQAQVLLGKRLKATL 660  
Qy 541 ELSPLVYSDKKKMSPPFICEBTEDEQLQTLIDGSNLEKELTNSRLKCRSDSPSPKK 600  
Db 661 ELSPLVYSDKKKMSPPFICEBTEDEQLQTLIDGSNLEKELTNSRLKCRSDSPSPKK 720  
Qy 601 TSLSPTSKLGSYSRDLDAKKKASLRQTESPPDADRTLTNADHSSKIVQRLRLSROE 660  
Db 721 TSLSPTSKLGSYSRDLDAKKKASLRQTESPPDADRTLTNADHSSKIVQRLRLSROE 780  
Qy 661 ELKERARVLLLEQARRDALAKGNKNTNTATPCNRLQSLQOQDEERRRQURERARQLIAB 720  
Db 781 ELKERARVLLLEQARRDALAKGNKNTNTATPCNRLQSLQOQDEERRRQURERARQLIAB 840  
Qy 721 ARSGVKKSELPSYGEAAETLKRSSKASG----- 780  
Db 841 ARSGVKKSELPSYGEAAETLKRSSKASG----- 865  
Qy 781 DLDTPQNSKLVDLKLKCLLEVOPOVANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQK 840  
Db 870 -----EQNSKLVDLKLKCLLEVOPOVANSPPSAQAQAVTSSSQDMKSGTEDLRTERLQK 924

```

Qy 841 TTERFRPVVFSKDSYTRKTQLOSPSOYIENRPMKORSIOBDYKKGNBKAITEOR 900
Db 925 TTERFRPVVFSKDSYTRKTQLOSPSOYIENRPMKORSIOBDYKKGNBKAITEOR 984
Qy 901 KPSEDEVYANKGFQDTSQYVVGELAALENBOKQIDTRALVEKRLRYLMDTGRNTEEBEAM 960
Db 985 KPSEDEVYANKGFQDTSQYVVGELAALENBOKQIDTRALVEKRLRYLMDTGRNTEEBEAM 1044
Qy 961 MOEWPMVNNKKNLIRRMNQLSLLEKSHDLERRYELNRELRAMLALIEDMQTEAQGRRE 1020
Db 1045 MOEWPMVNNKKNLIRRMNQLSLLEKSHDLERRYELNRELRAMLALIEDMQTEAQGRRE 1104
Qy 1021 QLLHDELVALVNRKDALVRDLDAOEKOABEEDHLEERTLEONKGMAKKEKCVLQ 1076
Db 1105 QLLHDELVALVNRKDALVRDLDAOEKOABEEDHLEERTLEONKGMAKKEKCVLQ 1160

RESULT 4
094977 HUMAN PRELIMINARY; PRT; 962 AA.
ID 094977_HUMAN
AC 094977_
DT 01-MAY-1999, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, sequence version 1.
DE KIAA0903 protein (Fragment).
GN Name=KIAA0903;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini; Homiidae;
OC Homo.
OC NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Brain;
RA MEDLINE=93156230; PubMed=10048485; DOI=10.1093/dnares/5.6.355;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirosewa M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RA Prediction of the coding sequences of unidentified human genes. XII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.
RL DNA Res. 5:355-364(1998).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative license
CC -----
DR EMBL; AB020710; BAA74926.1; -; mRNA.
DR HSSP; Q01082; 1BKX.
DR Ensembl; ENSG00000115504; Homo sapiens.
DR InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PSS0021; CH; 1.
DR NON TER
SQ SEQUENCE 962 AA; 109413 MW; 588BC31SD322D7AE CRC64;

Query Match 89.6%; Score 4919; DB 2; Length 962;
Best Local Similarity 100.0%; Pred. No. 1e-184;
Matches 962; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 181 LVWCKEVTKNYRGVKTITNFTTSWRNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAYD 240
Qy 355 GFSASIGISRLLESDVNLALIPDKLTWMTYLYOIRAHFSGQELANVOIENSSKSTYKVG 414
Db 241 GFASIGISRLLESDVNLALIPDKLTWMTYLYOIRAHFSGQELANVOIENSSKSTYKVG 300
Qy 415 NYETDINSSVDQEKFAELSDLKREPELOPISGAVDFLSQDSDSVFVNSGVESSEHQ 474
Db 301 NYETDINSSVDQEKFAELSDLKREPELOPISGAVDFLSQDSDSVFVNSGVESSEHQ 360
Qy 475 TPDDHLSPSTASPYCRRKSDTEPPQKSQSSGRTSSDDPGICSNSTDSYQAVLLGKRL 534
Db 361 TPDDHLSPSTASPYCRRKSDTEPPQKSQSSGRTSSDDPGICSNSTDSYQAVLLGKRL 420
Qy 535 LKAEFTLESLDYSDKKQDMSPPFCEETDEQQLQTLIDGSNLEKLENSRLBGRSDP 594
Db 421 LKAEFTLESLDYSDKKQDMSPPFCEETDEQQLQTLIDGSNLEKLENSRLBGRSDP 480
Qy 595 ESPFKKTSLSPTSKLGYSYRDLAKKQHASLRQTESPPDADRTTLNHAHSSKIYQHR 654
Db 481 ESPFKKTSLSPTSKLGYSYRDLAKKQHASLRQTESPPDADRTTLNHAHSSKIYQHR 540
Qy 655 LLSRQSELKERARVLLQARRDALAKAGKNTNTATPCNRQLSDQDEERRRLREPA 714
Db 541 LLSRQSELKERARVLLQARRDALAKAGKNTNTATPCNRQLSDQDEERRRLREPA 600
Qy 715 RQLIARBSGVKMSLEPSYGEAAEFLKERSKASGSENNIETDTEBELPEGVVGGGGE 774
Db 601 RQLIARBSGVKMSLEPSYGEAAEFLKERSKASGSENNIETDTEBELPEGVVGGGGE 660
Qy 775 LTNLSENDLDTPEONSGLVDLKLKLEVOPOVANSPPSAOKAVATSSSEODMKSGTEDLR 834
Db 661 LTNLSENDLDTPEONSGLVDLKLKLEVOPOVANSPPSAOKAVATSSSEODMKSGTEDLR 720
Qy 835 TTERLQTTFRFRNPVVFSSKDSYTRKTQLOSPSOYIENRPMKORSIOBDYKKGNBKA 894
Db 721 TTERLQTTFRFRNPVVFSSKDSYTRKTQLOSPSOYIENRPMKORSIOBDYKKGNBKA 780
Qy 895 ITTQCKPSEDEVYANKGFQDTSQYVVGELAALENBOKQIDTRALVEKRLRYLMDTGRNT 954
Db 781 ITTQCKPSEDEVYANKGFQDTSQYVVGELAALENBOKQIDTRALVEKRLRYLMDTGRNT 840
Qy 955 EEEERAMQEWPMVNNKKNLIRRMNQLSLLEKSHDLERRYELNRELRAMLALIEDMQTE 1014
Db 841 EEEERAMQEWPMVNNKKNLIRRMNQLSLLEKSHDLERRYELNRELRAMLALIEDMQTE 900
Qy 1015 AQKRREQLLDELVALVNRKDALVRDLDAOEKOABEEDHLEERTLEONKGMAKKEKCV 1074
Db 901 AQKRREQLLDELVALVNRKDALVRDLDAOEKOABEEDHLEERTLEONKGMAKKEKCV 960
Qy 1075 LQ 1076
Db 961 LQ 962

RESULT 5
0692M3 MOUSE PRELIMINARY; PRT; 1242 AA.
ID 0692M3_MOUSE
AC 0692M3_
DT 13-SEP-2004, integrated into UniProtKB/TrEMBL.
DT 13-SEP-2004, sequence version 1.
DT 07-FEB-2006, entry version 11.
DE KIAA0903 protein (Fragment).
GN Name=hnp1; Synonym=mk1A0903;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Fetal brain;

```



Rx PubMed:15368895; DOI=10.1093/dnares/11.3.205;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hirooka S.,  
 RA Suga Y., Seino S., Nishimura M., Kaiho T., Hoshino K., Kitamura H.,  
 RA Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT IV. The complete nucleotide sequences of 500 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries.";  
 RL DNA Res. 11:205-218(2004).  
 CC -----  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonCommercial license  
 CC -----  
 DR EMBL; AK173055; BAD32333.1; -; mRNA.  
 DR Ensembl; ENSMUSG0000042302; Mus musculus.  
 DR MGI; MGI:2667252; Ehhpl.  
 DR GO; GO:0005515; F:protein binding, IPI.  
 DR InterPro; IPR001715; Calponin\_act\_bd.  
 DR Pfam; PF00307; CH; 1.  
 DR SMART; SM00033; CH; 1.  
 DR PROSITE; PS50021; CH; 1.  
 DR NON TER 1  
 SQ SEQUENCE 1242 AA; 140356 MW; 1C8D17FC63BC0932 CRC64;  
 Query Match 85.3%; Score 4683.5; DB 2; Length 1242;  
 Best Local Similarity 84.1%; Pred. No. 2.2e-175;  
 Matches 945; Conservative 40; Mismatches 78; Indels 61; Gaps 10;  
 QY 1 MKQYASMPPTQTDVYKLFKPKLSKVVSAALQPSLSCTFLREGKATDEDMQSLASIVSMKQ 60  
 DB 132 MKQYASMPPTQTDVYKLFKPKLSKVVSAATLQPSLSCTFLREGKATDEDMQSLASIVSMKQ 191  
 QY 61 ADIGNLDFEEDNEDDDENRNVQBEKAKIT----- 91  
 DB 192 ADIGNLDFEEDNEDDDENRNVQBEKAKITELIVQNALSLDEDDQDCIKQANVPSAK 251  
 QY 92 -----ELINKLNFLEBEKDLATVNSNPFDDPAELINPGDPDSEETETASPRKTE 145  
 DB 252 SASSEBELINTLNLEDEKDLATVNTNPFDEPDTLEINPGDPDSEETETASPRKTE 311  
 QY 146 DSFTNNSYNPPEKAVQTPQYLNPFDEPEAVTIKDSPPQSTRKKNIRPVMSKYLVDSSK 205  
 DB 312 BSFTNNSCNPEKGVQTPQYLNPFDEPEAVTIKDSPPQSTRKKNIRPVMSKYLVDSSK 371  
 QY 206 TEEBELDESNPPEKSTPPPNNTLVNPOELSTERVYKKA-----PAPVLS 253  
 DB 372 S-BEELDESNPPEKPT-SPNNLVNTVQEGETERKVRKRAPAPAPAPAPAPAPALPT 429  
 QY 254 PKTGVLENT-VSAGKDLSTSPKPSPIPSVLAGRKPNASOSLIVWCKEVTNGVGVKTN 312  
 DB 430 PKTGV-NEVTVSAGKDLSTSPKPSPIPSVLAGRKPNASOSLIVWCKEVTNGVGVKTN 488  
 QY 313 FTTSWRNGLSFCALIHFRPDLIDYKSLNPQDIKENNKADGPAISIGISLSPSDVLT 372  
 DB 489 FTTSWRNGLSFCALIHFRPDLIDYKSLNPQDIKENNKADGPAISIGISLSPSDVLT 548  
 QY 373 LAIDKLTWMTYLYQIRAHFSGOELNVQIENSNSKSTYKGNTEITDNTSSVDQEKYAE 432  
 DB 549 LAIDKLTWMTYLYQIRAHFSGOELNVQIENSNSKSTYKGNTEITDNTSSVDQEKYAE 608  
 QY 433 LSLDKRBEPLQOPISGAVDFLSODDSVFVNDGSGESSEHOTDDHLSPTASPYCART 492  
 DB 609 LSLDKRBEPLQOPISGAVDFLSODDSVFVNDGSGESSEHOTDDHLSPTASPYCART 668  
 QY 493 KSDTEPQKSGSGSRTSGSDPGICSNITSTQAOVLGKKKLLKAELETLSLDLYVSDCK 552  
 DB 669 KSDTEPQKSGSGSRTSGSDPGICSNITSTQAOVLGKKKLLKAELETLSLDLYVSDCK 727  
 QY 553 DMSPPFCEETDEBQTLQTLIDGSLNLEKELNSRSLKGRSPSPSPKTSISLPSYKLGYS 612  
 DB 728 DVSP-----LSAYEBKTLQTVHASDDMEQGMKESRSLBECRLDGEIATITFNVSSPSKLG-- 781  
 QY 613 YSRRLDLAKKKGASLROTESPDPADRTTLNADHSKIVQRRLLSRQBELKERARVLLQ 672

DB 782 YNRDTEFTKFCASLRFOTESDPDADKTLTNADHPNKAQVHRMLSRQBELKERARVLLQ 841  
 QY 673 ARDPAALKAGKNTNTATPFCNRQLSDQDREBRRLRERARQLINAEAGVMSLPS 732  
 DB 842 ARDPAALFVSGKSGSAAPALCSRLNDQDDEBRRLRERARQLINAEAGVMSLPS 901  
 QY 733 YGMAAEKLEKRSKASGDENDINIDTNEIIPFEGFVVGCGDELNTLNDLDTPEONSLV 792  
 DB 902 YGMAAEKLEKRSKASGDENDINIDTNEIIPFEGFVVGCGDELNTLNDLDTPEONSLV 961  
 QY 793 DLKLLLEVPQVANSPPSSAQAQAVTSSSQDMKSGTEDLRTLRLOKTERFRNPVVS 852  
 DB 962 DLRLKLLLEVPQVANSPPSSAQAQAVTSSSQDMKSGTEDLRTLRLOKTERFRNPVVS 1021  
 QY 853 KDSIVRTQLOLFSQYIENRPEMRORSIQEDTKKGBEKAITETQKSPSEDTLVNGCF 912  
 DB 1022 KDSIVRTQLOLFSQYIENRPEMRORSIQEDTKKGBEKAITETQKSPSEDTLVNGCF 1078  
 QY 913 KDSIVYVGEIATLLENKOKIDTPALVEKRLRYLMDTGRVTEEBEAMQEMFVLVNGN 972  
 DB 1079 KDSIVYVGEIATLLENKOKIDTPALVEKRLRYLMDTGRVTEEBEAMQEMFVLVNGN 1138  
 QY 973 ALIRRNQSLLEKENDLERRYELINRELRLMLAIEDMQTEAKRRERQLLDELVALVN 1032  
 DB 1139 ALIRRNQSLLEKENDLERRYELINRELRLMLAIEDMQTEAKRRERQLLDELVALVN 1198  
 QY 1033 KRDLVVDLDAQKQASEDEHLERTLEONKGRKAKKEKVLQ 1076  
 DB 1199 KRDLVVDLDAQKQASEDEHLERTLEONKGRKAKKEKVLQ 1242  
 RESULT 6  
 Q5SQK3\_MOUSE PRELIMINARY; PRT; 1206 AA.  
 ID Q5SQK3\_MOUSE  
 AC Q5SQK3;  
 DT 21-DEC-2004, integrated into UniProtKB/TrEMBL.  
 DT 21-DEC-2004, sequence version 1.  
 DT 07-FEB-2006, entry version 10.  
 DE Novel protein.  
 GN Name=RP23-51G7.3; ORFNames=RP23-51G7.3-001;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridea; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Dunn M.;  
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RA Pelan S.;  
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.  
 CC  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution-NonCommercial license  
 CC -----  
 DR EMBL; AL731860; CA125363.1; -; Genomic DNA.  
 DR EMBL; AL669858; CA125363.1; JOINED; Genomic DNA.  
 DR EMBL; AL669858; CA125873.1; -; Genomic DNA.  
 DR EMBL; AL731860; CA125873.1; JOINED; Genomic DNA.  
 DR InterPro; IPR001715; Calponin\_act\_bd.  
 DR Pfam; PF00307; CH; 1.  
 DR SMART; SM00033; CH; 1.  
 DR PROSITE; PS50021; CH; 1.  
 SQ SEQUENCE 1206 AA; 136234 MW; D49618A5DD5D2A82 CRC64;  
 Query Match 82.9%; Score 4552; DB 2; Length 1206;  
 Best Local Similarity 82.3%; Pred. No. 3e-170;  
 Matches 925; Conservative 39; Mismatches 74; Indels 86; Gaps 11;  
 QY 1 MKQYASMPPTQTDVYKLFKPKLSKVVSAALQPSLSCTFLREGKATDEDMQSLASIVSMKQ 60

```
Db 121 MKQYASMPQTQDVKLFKFKPLSKKVSATLQFSLSCIFLRGKATDEDMQSLASLSMKQ 180
Qy 61 ADIGNLDDEEDNEDDENRVNOEKAKIT-----91
Db 181 ADIGNLDDEEDNEDDENRVNOEKAKITTEIVNQIALMSLDEDDDDCICQANVPSAK 240
Qy 92 -----ELINKNLFLEBAEKDLATVNSNPDDPAELNPGDPDSEEPITETASPRKTE 145
Db 241 SASSEBELINTLNFLEBAEKDLATVNTNPFDEPVTETELNPGDPDSE-----287
Qy 146 DSRFYNNSYNPFKEVQTYQYLNPFPEEPAPFTIKSPSPQSTKRKIRIVYMDSKITLYADSSK 205
Db 288 -----GVQTPQYLNPFDEPSTFWIKSPQSPQSTKRKIRIVYMDSKITLYADSSK 335
Qy 206 TEEBELDESNPFPYKSTPPNNLVNVOLETERRVKKA-----PAPVLS 253
Db 336 S-BEELDESNPFPYKSTPPNNLVNVOLETERRVKKA-----PAPVLS 393
Qy 254 PKTGVLNENT-VSAGKDLSTSPKSPISPVLGKRNASQSLVWCKEYTKYRGVKITN 312
Db 394 PKTGV-NENTVVSAGKDLSTSPKSPISPVLGKRNASQSLVWCKEYTKYRGVKITN 452
Qy 313 FTTSMWNGSLFCAILHFRPDLIDYKSLNPQDIKENNKAAYDGPASIGISRLLEPSDMVL 372
Db 453 FTTSMWNGSLFCAILHFRPDLIDYKSLNPQDIKENNKAAYDGPASIGISRLLEPSDMVL 512
Qy 373 LAIPDKLTWMTYLYOIRAHFSGOELNVVOIBENSSKSTYKGVNGETDINSVDOEKFYAE 432
Db 513 LAIPDKLTWMTYLYOIRAHFSGOELNVVOIBENSSKSTYKGVNGETDINSVDOEKFYAE 572
Qy 433 LSDLKEPELQPIGAVNDFLSQDDSVFVNDSGVGESESHQTPDDHLSPTASPYCRRT 492
Db 573 LSDLKEPEPHQARAGAVNDFLSQDDSVFVNDSGVGESESHQTPDDHLSPTASPYCRRT 632
Qy 493 KSTTEPKSQSSGRTSGSDPPIGCSNTDSTQAVLIGKRLKAEFTLSLTVSDLYSDPKK 552
Db 633 KSTTEPKSQSSGRTSGSDPPIGCSNTDSTQAVLIGKRLKAEFTLSLTVSDLYSDPKK 691
Qy 553 DMSPPFICEBTDEQIKOTLIGSNLEKKEKLENSRSLSCRSDDPESPIKSTLSPTSKLYGS 612
Db 692 DVSB---LSAYEOKLQTVHAASDMEGKMEKRSLECRLDGELAITKPNVSSPSKUG-- 745
Qy 613 YSRDLDAKKKHAHLRQTESDPADRTTLNADHSSKIVQHRLLSRQEBLKEBARVLLQ 672
Db 746 YNNDTPTTKPCASLRQIESDPADKSTLNHADHPNKAVOHRMLSRQEBLKEBARVLLQ 805
Qy 673 ARDDAALKAGNKANTNTATPFCNRQLSDQDEERRQLERARQLAEARSGVMSLEPS 732
Db 806 ARDDAALKAGNKANTNTATPFCNRQLSDQDEERRQLERARQLAEARSGVMSLEPS 865
Qy 733 YGMAAEKLEKRSKASDENDNT EIDTNEBI PGSPVVGDDDELTLNLENDTDEBONSIV 792
Db 866 YGMAAEKLEKRSKASDENDNT EIDTNEBI PGSPVVGDDDELTLNLENDTDEBONSIV 925
Qy 793 DLKALKLLEVOVANSPPSAQAQAVTESSEODMKSGTEDLARTERLOKTERFRNPVFS 852
Db 926 DLKALKLLEVOVANSPPSAQAQAVTESSEODMKSGTEDLARTERLOKTERFRNPVFS 985
Qy 853 KSTVTKTQLOSTFQYIENRPEMKGRSIOEDTKGNEERFAITETQRPKSEDEVANKGF 912
Db 986 KSTVTKTQLOSTFQYIENRPEMKGRSIOEDTKGNEERFAITETQRPKSEDEVANKGF 1042
Qy 913 KQTSQYVVGELALENOKQIDTRPALVEKRLRYLMDTGNTTEEBEEMQEMFMLVANKN 972
Db 1043 KQTSQYVVGELALENOKQIDTRPALVEKRLRYLMDTGNTTEEBEEMQEMFMLVANKN 1102
Qy 973 ALIRRNQSLSLKEHDLERREYELINRELRAMLAIEDMQTEBAQKREQLLDELVALVN 1032
Db 1103 ALIRRNQSLSLKEHDLERREYELINRELRAMLAIEDMQTEBAQKREQLLDELVALVN 1162
Qy 1033 KRDAVLDLDAOEKQAEDEDEHLERTLEONKGMKAKKEKCVLQ 1076
```

```
Db 1163 KRDAVLDLDAOEKQAEDEDEHLERTLEONKGMKAKKEKCVLQ 1206
RESULT 7
ID Q912J6_MOUSE PRELIMINARY; PRT; 1206 AA.
AC Q912J6;
DT 01-DEC-2001, integrated into UniProtKB/TrEMBL.
DT 01-DEC-2001, sequence version 1.
DT 07-FEB-2006, entry version 19.
DE KIAA0903-like protein.
GN Name=Bhdp1; Synonym=Af424697;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=C57BL/6J;
RX MEDLINE=2325413; PubMed=12174196; DOI=10.1186/1471-2156-3-14;
RA Ruch S., Resch K., Thiel C., Ulbrich M., Platzer M., Jockusch H.,
RA Schmitz-John T.;
RT "Comparative transcription map of the wobbler critical region on mouse
RT chromosome 11 and the homologous region on human chromosome 2p13-14.";
RL BMC Genet. 3:14-14(2002).
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs License
CC -----
DR EMBL; AF424697; AAL24806.1; -; mRNA.
DR HSSP; Q01082; 1BKR.
DR MGI; MGI:2667252; Bhdp1.
DR GO; GO:0005515; F-protein binding; IPI.
DR InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS50021; CH; 1.
SQ SEQUENCE 1206 AA; 136047 MW; 349A9A7C7CF0AC46 CRC64;

Query Match 82.6%; Score 4532; DB 2; Length 1206;
Best Local Similarity 82.0%; Pred. No. 1,8e-169;
Matches 922; Conservative 39; Mismatches 77; Indels 86; Gaps 11;

Qy 1 MKQYASMPQTQDVKLFKFKPLSKKVSATLQFSLSCIFLRGKATDEDMQSLASLSMKQ 60
Db 121 MKQYASMPQTQDVKLFKFKPLSKKVSATLQFSLSCIFLRGKATDEDMQSLASLSMKQ 180
Qy 61 ADIGNLDDEEDNEDDENRVNOEKAKIT-----91
Db 181 ADIGNLDDEEDNEDDENRVNOEKAKITTEIVNQIALMSLDEDDDDCICQANVPSAK 240
Qy 92 -----ELINKNLFLEBAEKDLATVNSNPDDPAELNPGDPDSEEPITETASPRKTE 145
Db 241 SASSEBELINTLNFLEBAEKDLATVNTNPFDEPVTETELNPGDPDSE-----287
Qy 146 DSRFYNNSYNPFKEVQTYQYLNPFPEEPAPFTIKSPSPQSTKRKIRIVYMDSKITLYADSSK 205
Db 288 -----GVQTPQYLNPFDEPSTFWIKSPQSPQSTKRKIRIVYMDSKITLYADSSK 335
Qy 206 TEEBELDESNPFPYKSTPPNNLVNVOLETERRVKKA-----PAPVLS 253
Db 336 S-BEELDESNPFPYKSTPPNNLVNVOLETERRVKKA-----PAPVLS 393
Qy 254 PKTGVLNENT-VSAGKDLSTSPKSPISPVLGKRNASQSLVWCKEYTKYRGVKITN 312
Db 394 PKTGV-NENTVVSAGKDLSTSPKSPISPVLGKRNASQSLVWCKEYTKYRGVKITN 452
Qy 313 FTTSMWNGSLFCAILHFRPDLIDYKSLNPQDIKENNKAAYDGPASIGISRLLEPSDMVL 372
Db 453 FTTSMWNGSLFCAILHFRPDLIDYKSLNPQDIKENNKAAYDGPASIGISRLLEPSDMVL 512
Qy 373 LAIPDKLTWMTYLYOIRAHFSGOELNVVOIBENSSKSTYKGVNGETDINSVDOEKFYAE 432
```

```

Db      513 LAIPKLTVMYTYLYQIRAHFSGGELNVVQIENSSKSTYKGVNGETDNTSSVDQKPYAE 572
Qy      433 LSDLRKEPELOQPIGAVDFLSQDDSVFVNDGSGVESESEHQTPDHLSPSPASRYCRT 492
Db      573 LSDLRKEPEPHOPANGAVDLISQDDSVFVNDGSGVESESEHQTPDHLSPSPASRYCRT 632
Qy      493 KSDTEPQKSGSGRTSGSDPGICSNSTDSQAVLGGKRLKKAETLSDLYVSDKKK 552
Db      633 KSDTEPQKSGSGRTSGSDPGICSNSTDSQAVLGGKRLKKAETLSDLYVSDKKK 691
Qy      553 DMSPPFIEETDEQQLQTLIDIGSNLEKELNSRLSECRSPESPISIKTSLSPISKLYS 612
Db      692 DVSP-----LSAYEQQLQTVHAASDMEQCKMEKRSLECRLEGLAITKPNVSPSKLG-- 745
Qy      613 YSRDLIDLAKKKHASTROTESPDPDARTLNADHSKTVQHRLLSROELKERAVLLEQ 672
Db      746 YNRDLDFPKKCAIRQIESPDADKSTLNADHPNKAVQHRMLSRQELKERAVLLEQ 805
Qy      673 ARDDAALKAGNKNHTNTATPCNROLSDQDEERRRQLERARQLIAARSGVKMSLPS 732
Db      806 ARDDAALKAGNKNHTNTATPCNROLSDQDEERRRQLERARQLIAARSGVKMSLPS 865
Qy      733 YGMAAERLKERKSKASGDENDNIEIDTNEBIEPGFVVGSGDELITLNDLTPGQNSKLV 792
Db      866 YGMAAERLKERKSKASGDENDNIEIDTNEBIEPGFVVGSGDELITLNDLTPGQNSKLV 925
Qy      793 DLKAKLLEVPQVANSFSSAAQKAVTSSSEODMSGSGEDLTERLOQTTERFRNPVFS 852
Db      926 DLKAKLLEVPQVANSFSSAAQKAVTSSSEODMSGSGEDLTERLOQTTERFRNPVFS 985
Qy      853 KSTVTRKTQLOSGFSGYIENRPEMKRORSIOEDTKKNGEKAITETOKSPSEDTLNGKF 912
Db      986 KSTVTRKTQLOSGFSGYIENRPEMKRORSIOEDTKKNGEKAITETOKSPSEDTLNGKF 1042
Qy      913 KDTSGYVVGELALLENQKQIDTRALVVEKRLRYLMDTGRNTEBEANMOEFMLVNNKN 972
Db      1043 KDTSGYVVGELALLENQKQIDTRALVVEKRLRYLMDTGRNTEBEANMOEFMLVNNKN 1102
Qy      973 ALIRRMQSLLEKSHDLERREYELNRLRLMALIEDMOQTBAQRRQLLDLVALVN 1032
Db      1103 ALIRRMQSLLEKSHDLERREYELNRLRLMALIEDMOQTBAQRRQLLDLVALVN 1162
Qy      1033 KRDLAVRDLDAQKQAEDEDEHLEKTLLEQNGKMAKKEKCVLQ 1076
Db      1163 KRDLAVRDLDAQKQAEDEDEHLEKTLLEQNGKMAKKEKCVLQ 1206

```

## RESULT 8

```

053TV6 HUMAN PRELIMINARY; PRT; 893 AA.
AC Q31V6;
DT 24-MAY-2005, integrated into UniProtKB/TrEMBL.
DT 07-FEB-2006, entry version 3.
DE Hypothetical protein KIAA0903 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA "The sequence of Homo sapiens BAC clone RP11-443F16."
RT Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.
RL NUCLEOTIDE SEQUENCE.
RA Waterston R.H.;
RL Submitted (OCT-2000) to the EMBL/Genbank/DBJ databases.
RP NUCLEOTIDE SEQUENCE.

```

```

RA Waterston R.H.;
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP NUCLEOTIDE SEQUENCE.
RA Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/Genbank/DBJ databases.
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivs license
CC -----
DR EMBL; AC007098; AA14789.1; -; Genomic DNA.
DR Ensembl; ENSG00000115504; Homo sapiens.
DR InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PS50021; CH; 1.
DR Hypothetical protein.
FT NON_TER 1
FT TER 893
SQ SEQUENCE 893 AA; 100446 MW; 88B387A63176A52E CRC64;

Query Match      80.4%; Score 4414.5; DB 2; Length 893;
Best Local Similarity 97.5%; Pred. No. 5.4e-165;
Matches 864; Conservative 6; Mismatches 15; Indels 1; Gaps 1;

Qy      66 LDDPEEDND-DDENRVNGEKAKITELINKLFLDEAEKDLATVNSNPDPDAEELN 124
Db      8 LSSIDEDDDDCIKANNMSAKSASSSEELINKLFLDEAEKDLATVNSNPDPDAEELN 67
Qy      125 PFGDPDESEPTTETASPRKTEDSFYNNSYNPFKEVQTPQYLNPPDEPEAFYTIKDSPPQS 184
Db      68 PFGDPDESEPTTETASPRKTEDSFYNNSYNPFKEVQTPQYLNPPDEPEAFYTIKDSPPQS 127
Qy      185 TKRNIRPVDMSKLYVADSSKTEBEELDESNPYEKSTPPNNLVNPFQELFETRRYKR 244
Db      128 TKRNIRPVDMSKLYVADSSKTEBEELDESNPFEKSTPPNNLVNPFQELFETRRYKR 187
Qy      245 KAPAPVLSPTGYLNNNTVYAGKDLSTSPKPSPIPSPVGRKXNASQSLVWCKEYTKN 304
Db      188 KAPAPVLSPTGYLNNNTVYAGKDLSTSPKPSPIPSPVGRKXNASQSLVWCKEYTKN 247
Qy      305 YRGVKTINFTTWSNGSLFCALHFRPDLIDYKSLNPQDIKENNKAKAYDFASIGISRL 364
Db      248 YRGVKTINFTTWSNGSLFCALHFRPDLIDYKSLNPQDIKENNKAKAYDFASIGISRL 307
Qy      365 LEPSDMVLAI PDKLTVMTYLYQIRAHFSGGELNVVQIENSSKSTYKGVNGETDNTSSV 424
Db      308 LEPSDMVLAI PDKLTVMTYLYQIRAHFSGGELNVVQIENSSKSTYKGVNGETDNTSSV 367
Qy      425 DQKFTYAEISDLKREPELOQPIGAVDFLSQDDSVFVNDGSGVESESEHQTPDHLSPST 484
Db      368 DQKFTYAEISDLKREPELOQPIGAVDFLSQDDSVFVNDGSGVESESEHQTPDHLSPST 427
Qy      485 ASPYCRRTKSDTEPQKSGSGRTSGSDPGICSNSTDSQAVLGGKRLKKAETLSD 544
Db      428 ASPYCRRTKSDTEPQKSGSGRTSGSDPGICSNSTDSQAVLGGKRLKKAETLSD 487
Qy      545 LYSVDKKKMSPPFICEETDEQQLQTLIDIGSNLEKELNSRLSECRSPESPISIKTSLS 604
Db      488 LYSVDKKKMSPPFICEETDEQQLQTLIDIGSNLEKELNSRLSECRSPESPISIKTSLS 547
Qy      605 PTKSLGYSYSRDLIDLAKKKHASTROTESPDPDARTLNADHSKTVQHRLLSROELKE 664
Db      548 PTKSLGYSYSRDLIDLAKKKHASTROTESPDPDARTLNADHSKTVQHRLLSROELKE 607
Qy      665 RARVLEQARRDALKAGNKNHTNTATPCNROLSDQDEERRRQLERARQLIAARSG 724
Db      608 RARVLEQARRDALKAGNKNHTNTATPCNROLSDQDEERRRQLERARQLIAARSG 667
Qy      725 VKMSLEPSYGEAAERLKERKSKASGDENDNIEIDTNEBIEPGFVVGSGDELITLNDLDT 784
Db      668 VKMSLEPSYGEAAERLKERKSKASGDENDNIEIDTNEBIEPGFVVGSGDELITLNDLDT 727

```

```

QY 785 PEONSKLVLDLKKLLLEVOPOVANS PSSAOKAVTESSEODMKGTEDLRTERLQKTTTER 844
DB 728 PEONSKLVLDLKKLLLEVOPOVANS PSSAOKAVTESSEODMKGTEDLRTERLQKTTTER 787
QY 845 FRNPVFSKSTVAKTQLOQSFQYIENRPEMKORSIOEDTKGNEKAATTEORPSE 904
DB 788 FRNPVFSKSTVAKTQLOQSFQYIENRPEMKORSIOEDTKGNEKAATTEORPSE 847
QY 905 DEVLANKFPDTSQYVVGELALALENEKQIDTRALVERKRLYLMDT 950
DB 848 DEVLANKFPDTSQYVVGELALALENEKQIDTRALVERKRLYLMDT 893

RESULT 9
ID Q4SG89_TETNG PRELIMINARY; PRT; 1036 AA.
AC Q4SG89;
DT 19-JUL-2005, integrated into UniProtKB/TrEMBL.
DT 19-JUL-2005, sequence version 1.
DE 07-FEB-2006, entry version 4.
DE Chromosome 17 SCF14597, whole genome shotgun sequence.
GN ORFNames=GSTENG0018711001;
OS Tetradodon nigroviridis (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygii; Acanthopterygii; Percormorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_taxid=99883;
RN Nucleotide Sequence.
RX PubMed=15496914; DOI=10.1038/nature03025;
RA Jallion O., Aury J.-M., Brunet F., Petit J.-L., Strange-Thomann N.,
RA Mucelli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Nicard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Daelva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Anthouard V., Ublin C., Castellani V., Katinka M., Vacherie B.,
RA Blemont C., Skalli Z., Cattoiello L., Poulin J., De Berardinis V.,
RA Cuend C., Duprat S., Broctier P., Couanceau J.-P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellis M., Wolf J.-N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Lander V., Schachter V., Queller F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.;
RA "genome duplication in the teleost fish Tetradodon nigroviridis reveals
RT the early vertebrate proto-karyotype.";
RT Nature 431:946-957(2004).
RN Nature 431:946-957(2004).
RP Nucleotide Sequence.
RG Genoscope, Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/Genbank/DBJ databases.
CC -1- CATION: The sequence shown here is derived from an
CC EMBL/Genbank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
CC -----
CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms
CC Distributed under the Creative Commons Attribution-NonDerivative license
CC -----
CC EMBL, CAAB01014597; CAG00313.1; -; Genomic DNA.
CC InterPro; IPR001715; Calponin_act_bd.
DR Pfam; PF00307; CH; 1.
DR SMART; SM00033; CH; 1.
DR PROSITE; PSS0021; CH; 1.
SQ PROQUENCE 1036 AA; 114819 MW; CLA8E829F17262B6 CRC64;

Query Match 41.3%; Score 2265.5; DB 2; Length 1036;
Best Local Similarity 49.1%; Pred. No. 9.1e-81;
Matches 524; Conservative 121; Mismatches 188; Indels 235; Gaps 30;
QY 1 MKQYASMPPTQTVKTKFKPLSKRVVSAALQFSLCTFLRSGKATDDEDMQSLAVSMKQ 60
DB 32 MKQYASMPPTQTVKTKFKPLSKRVVSAALQFSLCTFLRSGKATDDEDMQSLAVSMKQ 91
QY 61 ADGNLDDPFEBNEDDDENRNVNQEKAAKT-----ELINLNFLEDAEK 105

```

```

DB 92 ADGNLDDPFEBNEDDDENRNVNQEKAAKTIVFLISFLVCTFVALLEIKLNFLEDKQ 151
QY 106 DLN-TVNSNPDPDPA-----ELNPPDPPEEPITETASPKRTEDSFYNSTNPFK--- 157
DB 152 EASPMSTNPFDEPDDLHFNHFNPFEDPMEBAAPQPNRSRLDSDNDSNPFEPD 211
QY 158 --EYQFPQYLNPFDEPEAFVTK--DSPPQSTRKKNIRPDMSKYLYADSSKTEBEELDE 213
DB 212 EPELQTPG--NPFDEPDQDTLDQDEPEFPKPRQKGVYVDMSKLYLADVNHDEDELD 269
QY 214 -----SNPFYPPKSTPPNNLVNPFQELETERRYKKAAP 249
DB 270 FTQLPLGCVNERSMNCHFCILRSNPFYEPKTSFPAKPVGQ--PSLDMSSQKRAPPP 328
QY 250 PVLSPKTV-----VLNNTYSA--GDL-STSP-----K 275
DB 329 PSSSPGFPSPAPSKSSVDRBOALAVGSPVTAVALGRELASSSPKMLQGIATIVSLARK 388
QY 276 PSPSPVLGRKPNASQSLVWCKEYTKNRYGKTNFTTSMRNGLSFCALHFRPDLI 335
DB 389 PSPSPVLGRKPNASQSLVWCKEYTKNRYGKTNFTTSMRNGLSFCALHFRPDLI 448
QY 336 DYKSLNPQDIKENNKAYDGFASIGISRLLEPSDVLALPDKLTVWTYLYQIRAFPSGQ 395
DB 449 DYKSLNPQDIKENNKAYDGFASIGISRLLEPSDVLALPDKLTVWTYLYQIRAFPSGE 508
QY 396 ELNVVQIEENSSKSTYKNGYFETDNTSSVVDKFCYAEISLQKPELQGI---SGAVDF 452
DB 509 ELNVVQIEENSSKSTYKNGYFETDNTSSVVDKFCYAEISLQKPELQGI---SGAVDF 568
QY 453 LSQDSDV-----FVNSGVSESESEHQTDPDHLSPSTASPYCARTKSDTEPQKQSSGR 507
DB 569 TKNEDEGPRKADCNVDKAVLTKTGHO-----QSSPVPSP---RTGVYTVSDSSQA--- 618
QY 508 TSGSDDPGICSNSTDSQOAVLKKLLKAELELSDLVYSDKKKMSPEFICEETDEQK 567
DB 619 TPASBERG-----SLTKANTLDSLPRERB----- 645
QY 568 LQTLIDGSLNLEKLENSR-----SLGCRSDPSPPIKKTSLSP---SKLGYSRDLDA 620
DB 646 -----MEKKQOQKGVGEETNENGTPEGSPGRGPLSSPHHQKLGFSYNRDLDI 696
QY 621 KKKHSLRQTESPDPRDRT--TLNHADH-----SSKLVQHRLLSROBELKERAR 667
DB 697 KKKHSLRQTESPDPRDRT--TLNHADH-----SSKLVQHRLLSROBELKERAR 756
QY 668 VLEQARRDALAKGNKNTNTATPPCN--RQSDQDEBERRQLRERARQLIAEASGV 725
DB 757 LLEQARRDALAKGNKNTNTATPPCN--RQSDQDEBERRQLRERARQLIAEASGV 816
QY 726 KMELELSYGMAAEKLL--EKSASGDEBNIEIDTBEIPEGVVGSGDELTYLENDLDT 784
DB 817 KMELELSYGMAAEKLL--EKSASGDEBNIEIDTBEIPEGVVGSGDELTYLENDLDT 857
QY 785 PEONSKLVLDLKKLLLEVOPOVANS PSSAOKAVTESSEODMKGTEDLRTERLQKTTTER 844
DB 858 -----SERGIGRAITAGSHR----- 873
QY 845 FRNPVFSKSTVAKTQLOQSFQYIENRPEMKORSIOEDTKGNEKAATTEORPSE 903
DB 874 FRNPVFSKSTVAKTQLOQSFQYIENRPEMKORSIOEDTKGNEKAATTEORPSE 933
QY 904 EDEVLANKFPDTSQYVVGELALALENEKQIDTRALVERKRLYLMDT 951
DB 934 EDEVLANKFPDTSQYVVGELALALENEKQIDTRALVERKRLYLMDT 981

RESULT 10
ID Q5POM3_RAT PRELIMINARY; PRT; 762 AA.
AC Q5POM3;
DT 04-JAN-2005, integrated into UniProtKB/TrEMBL.

```



RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=BALB/c;  
 RA Melichar J.M., Noegel A.A., Korenbaum E.;  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RA MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marzella K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carrincci P., Prange C.,  
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulvik S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Gilwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;  
 RA Straube R.;  
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
 CC Copyrighted by the UniProt Consortium, see http://www.uniprot.org/terms  
 CC Distributed under the Creative Commons Attribution-NonDerivs License  
 CC  
 DR EMBL: AF305088; AAK32110.1; -; mRNA.  
 DR EMBL: BC004660; AA044660.1; -; mRNA.  
 DR HSSP: 001082; IBKR.  
 DR MGI: MGI:1341841; Kcnk7.  
 DR GO: GO:0016021; C:Integral to membrane; RCA.  
 DR GO: GO:0016020; C:membrane; RCA.  
 DR GO: GO:0005267; F:Potassium channel activity; RCA.  
 DR GO: GO:0006812; P:cation transport; RCA.  
 DR GO: GO:0006813; P:potassium ion transport; RCA.  
 DR InterPro: IPR001715; Calponin\_act\_ba.  
 DR Pfam: PF00307; CH; 1.  
 DR SMART: SM00033; CH; 1.  
 DR PROSITE: PS50021; CH; 1.  
 SQ SEQUENCE 874 AA; 95561 MW; 94803958BCDF5B6D CRC64;  
 Query Match 21.0%; Score 1151.5; DB 2; Length 874;  
 Best Local Similarity 30.3%; Pred. No. 3.2e-37;  
 Matches 337; Conservative 116; Mismatches 264; Indels 395; Gaps 27;  
 QY 1 MKOYASMPPTQDTKLFKPLSKKVVSAALQFSLCLFLRGKATDDEMOGLASVGMQ 60  
 DB 120 LAHAGGVPAQVPLKRLKPSKVVVAHSLTSLGVLBRGRTDMDQSLASLMSVKP 179  
 QY 61 ADIGNLDDPEEDNEDDENRNOEKA-AKITEILINKLNPIDEAKDLATVNSNPFDDPD 119  
 DB 180 SDVGNLDDFAESBEEBANGAPAVRTGRQSDLSRLKTLCE-EBDEGHIR-----PQ 232  
 QY 120 AAEINPFGDDP-SSEPTETASPKTSDSYNNSYNPFKEVQTPQVYINPDEPAFYTIK 178  
 DB 233 QAAARSSASADTSPAPVSAAPAPVRA-----FRGQSSBPALITGCG 273  
 QY 179 DSPPGQTKRKINIRVDMSKVLVADSSKTBEEELDESFPYEPKTPPPNNLVNPQLEET 238  
 DB 274 VGP-----ETPPBPSPSP-ETSTGTGPGGTWFTP--AP 304  
 QY 239 ERRVKKAPAPV-----LSPKGTVLNENTVSA-----GDL-STSPPKSPIP 280

DB 305 RLKRGSDAPSSPVPCSGDEVPTNSDEPPTGMSSGGERQAQISQEGTEAHEARPEPIEA 364  
 QY 281 SPVLGRKRNASOSLILVCKEYVATKYRGVKTNTTTSWRNGLSRCAIHHRRPLDIYKSL 340  
 DB 365 PPTL---VSSOSLLEWCQEVINGRGVCTINFTTSWRNGLARCAIHHRRYPDKIDYFSL 421  
 QY 341 NPDIKENNKKAADGPAFISIGRLBPSDMVLLAIPDKLTVMYLYVQIRAHFSGQELNV 400  
 DB 422 DPLAIKNNKQADGPAALGVSLBEPADNVLLSVPKLIVMYTLLQIRAFCTGQELQV 481  
 QY 401 QIEENSSKSTYKGVNVTDTNNSVDQEKFYAELSDLKREBELQOPT-----SGAVD 451  
 DB 482 QLEGGGSGGYRVGNAPSLPDCLDA---GDIAQLRLEHGAIVPEPKAVNRGTGAIP 537  
 QY 452 FLEQDDSVFVNDGVSSESSEHQTDPDHLSPSTASPCRTKSDTEBQKQSGSRSSG 511  
 DB 538 KVASRPTDLSCKSRDGALE-----VAQBAIPQBAPTDGPARRSS 576  
 QY 512 DDP-----GICSTDTSTQAQVLLGKRLKAKTLBLSDLVVSXKKDMSPPICEETDEQ 566  
 DB 577 TTIVVPAEGLVNGVGS-----GVRLRRS----- 601  
 QY 567 KLTGLDIGNLKEKLENSRLSECRSDPESPIKTSLSPTSGLGYSRDLAKKXHAS 626  
 DB 602 -----SYNGBAQVPP-----PRAHGSFQHVADADLKKRRSR 634  
 QY 627 LRQTSDDPADRTTLTHADHSKIYQHRLLSRBELKERARVLLQARRDALAKGNKN 686  
 DB 635 LR-----N 637  
 QY 687 TMTATPFCNQRQLSDODDEBRRLQRRBARQLIABRGVMSLPSYGAARKEKRSK 746  
 DB 638 SNRPS-----VDDQD-----SGAAVAGPA-GPGAVP----- 663  
 QY 747 ASGDENDNIETDNEBIPFGFVVGGBDELTNLENDLDTPEONSLVDLKLKLLVQPVQ 806  
 DB 664 -----GPNP 667  
 QY 807 ANPSSAAQAATVESSEBOMKSGTEDLRTERLOKTERFRANPVVFSKOSTVTRKTQLSFS 866  
 DB 668 ASSPDNPPLPAPVPOQPPGPPPT----- 692  
 QY 867 QYLENPEMKORSIEDTQKKEEKAITETQRKSEDEVLNKGFPTDSQYVVGELAL 926  
 DB 693 ---ESSPSLGEELGRLR-----FQDTSQYCAELQAL 721  
 QY 927 ENEQKIDTRALVERKRLRYLMDTGNTEBEEAMQEWFLVYKKNALIRRNQLSLK 986  
 DB 722 EQGQGIDGRAAEVKEQLRSLMESGANRLOEBVLQEMFTLVKKNALIRRODQLDLIE 781  
 QY 987 EHLERRRYELNRELDAALAIEDMQTEAKRREQLLDELVALVYKKNALVDDDAQEK 1046  
 DB 782 EQLERRFELSLRELAMLAIEEMQKTVQAQHQEQLLBELVSLVQORDELVDLDQKER 841  
 QY 1047 QAEEDHELERTLQONKGMKAK---KEBKVLL 1075  
 DB 842 IALEEDERLRLRGLEQRKRVSRQLSRRECTL 873  
 RESULT 12  
 Q99MS6\_MOUSE PRELIMINARY; PRT; 759 AA.  
 AC Q99MS6;  
 DT 01-JUN-2001, integrated into UniProtKB/TrEMBL.  
 DT 07-FEB-2006, entry version 17.  
 DE Tangerin C.  
 GN Name=Kcnk7;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
 OC Muridae; Murinae; Mus.





Matches 327; Conservative 108; Mismatches 195; Indels 466; Gaps 26;

QY	1	WKQVYASPMPTQTDVTKKREPKLSKVYSAALQPSLSCITPLREBKATDDEMOIASIVSMKQ	60
Db	120	LAHHAGVPVPAQVPLRLRKPKSVKVVHVELSLTISGLVILREGRATDDMOIASIAMSVPK	179
QY	61	ADIGLUDPPEENEDDENRNVQEEKAKITELINKLNFLEBAEKDLATVNSNPPDDPA	120
Db	180	SDVGNLDDPABESDE-----	194
QY	121	AELNPFQDPDSEEPITETASPKTEDSFYNNSYNPKEVQTPQYLNPDPDEAPVTLIKDS	180
Db	195	-EANGPAPAE-----	207
QY	181	PPQSTK---RKNIIPVDMSKTYLVDSSKTEBEELEDSNPFPYEPKSTPPNNLVNPFVEL	236
Db	208	GFQSGRGCAFRJGRPP-DLSREL-----KTLCEEBEDEGH-----IRPQO--	245
QY	237	ETERYVKKAPRPPYLSPTGYLANENTVYSAQDLSTSPKPSPIPSVLGRKPMASQSLV	296
Db	246	-----AARP-----SSAED--TSPAPAP--PTL--VSSQSLLLE	274
QY	297	WCKEYTKNVYRGVYITNFTTSMRNGSLPCALILHFRPDDILDYSLNPODIKENNKAYAGGF	356
Db	215	WQGEYTNRGVCYCTNFTTSMRNGALFCAILHFRFPDIDFSLDPLNIKQNNQAPFGCF	334
QY	357	ASIGISRLLEPSDWLLAI PDKLTVMTYLYQIIRAHFSQGEILNVVQIIEENSKSTYKVNY	416
Db	335	AALGVSRLLERPADVYLVSVPDCLIMTYLTCQIRAFCTGQELQVLQLEGGGSGGTYRVGNA	394
QY	417	ETDNTSSVDQEKFYALELSDLKREPLQOPI-----SCAVDPLSDDDSVFVNDSGVG	467
Db	395	QPSLTPDCIDA---GDLAQRLREHGAEPTEBEKAVNNGTGAIPVVASRDITDLSGSSXDG	450
QY	468	ESESEHQTPDDHLSPSTASPYCRKFKSTBPQKQSGSGTSGSDDP-----GLICSNDS	522
Db	451	EAE-----VAQBAI PQEALTPDPBARSSITPPVPALEGVLNNGGA	489
QY	533	TQAOYLLKQKLLKAETLELSDLYVSDKKQMSPPFICEETDEQKLTLDIGSNLEKEKL	582
Db	490	S-----GAVRLRRS-----	498
QY	593	ENSRLEGRSDPEESIKTISLSPYSKLGYSYSRDLDAKKGHASLROTESPDDADRTTLN	642
Db	499	---SVNGEAGVPP-----PRAHGSFISHRADLLKGRSRLR-----	533
QY	643	HADHSKIVQRIHLSRQELKERBAVLLLEQARBAALKAQNKGNNTATPFCNRQLSPDQ	702
Db	534	-----NSNGTS-----VDDQ	543
QY	703	DEERRRLRERARQULIARSGVKNSELPSYGENAAEKLKERSKASGDENDIBIDTNEE	762
Db	544	D-----SGAAVGAQPA-GPAVE-----	560
QY	763	IPBGFVVGCGDELTNLENDLDTPEQNSKLVLDLKLKGLLEVPQVANSPPSSAAQAVTSS	822
Db	561	-----GPNZASSPDANPLPAVPQO	580
QY	823	BQDMKSGTEDLRTERLOKTTFRFNPVVPFSKDSYTRKTQLOSPQYIENRBMKRSIQ	882
Db	581	PPGGEPPTE-----BSSSLBEGGLQ	602
QY	883	EDTKKGNNEKAITETQRPSEDEVLANKGFKDTSQYVVGELALALENEKOQIDTAAALVEK	942
Db	603	R-----FQDTISGYCAELQALEBEGQIDGAAAVEX	634
QY	943	RLRYLMDTGRNTEEBEAMQEMFLVNKKALIRBNQLSLBKXHDJERRYYELLNREL	1002
Db	635	QLRSIMESGARLTOBEVLIQEMFTLVNKKNALIRQDQLOLLIEBODLERRFELLSREL	694
QY	1003	AMLAIEDMOCKTEAQGREROLLDELVALVNRKDALVRLDLQAEQAEDEDHLERTLEON	1062
Db	695	AMLAIEBMOCKTVAAQHREROLLLEELVSLVNRDDELVRDLDERKARALEEDRERLGRLEQR	754

QY 1063 KGMKAK---KEEKCVL 1075  
ID 755 RRVKVSQLSRRERCTL 770

RESULT 14  
OEC2W6 MOUSE  
ID OEC2W6 MOUSE PRELIMINARY; RT; 723 AA.  
AC OEC2W6;  
DT 01-MAR-2003, integrated into UniProtKB/TrEMBL.  
DT 01-MAR-2003, sequence version 1.  
DT 07-FEB-2006, entry version 16.  
DE 2 days pregnant adult female ovary cDNA, RIKEN full-length enriched library, clone:R330024K12 product:cangerin, full insert sequence.  
DS (fragment).  
GN Name=Kcnk7;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;  
OC Muridea; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
NM NCULEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Ovary;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RT Carninci P., Hayashizaki Y.,  
RL "High-efficiency full-length cDNA cloning",  
RN Methods Enzymol. 303:19-44(1999).  
[2]  
NM NCULEOTIDE SEQUENCE.  
RC STRAIN=C57BL/6J; TISSUE=Ovary;  
RX PubMed=16141072; DOI=10.1126/science.1112014;  
RT Carninci P., Katukawa T., Katayama S., Gough J., Frith M.C., Meda N.,  
RX Oyama R., Ravasi T., Lenhard B., Wells C., Kodzius R., Shimokawa K.,  
RA Bajic V.B., Bremner S.E., Batalov S., Forrest A.R., Zavolan M.,  
RA Davis M.J., Wilming L.G., Altadov S., Allen J.E.,  
RA Ambesi-Impombato A., Aweissel R., Atturliya R.N., Bailey T.L.,  
RA Banasi M., Baxter L., Beisel K.W., Bersano T., Bono H., Chalk A.M.,  
RA Chiu K.P., Choudhary V., Chircioffels A., Clutterbuck D.R.,  
RA Crowe M.L., Dalla E., Dalrymple B.P., De Bono B., Della Gatta G.,  
RA Di Bernardo D., Down T., Engstrom S.P., Fagiolini M., Faulner G.,  
RA Fletcher C.F., Fukushima T., Furuno M., Futaki S., Gariboldi M.,  
RA Geisler-Hemming P., Gingeras T.R., Gojobori T., Green R.B.,  
RA Guetlich S., Harbers M., Hayashi Y., Hensch T.K., Hirokawa N.,  
RA Hill D., Hummel L., Iacono M., Ikey K., Iwama A., Ishikawa T.,  
RA Jakt M., Kanapin A., Katoh M., Kawasawa Y., Kelso J., Kitamura H.,  
RA Kitano H., Kollias G., Krishnan S.P., Kruger A., Kummerfeld S.K.,  
RA Kurochkin I.V., Lacroix L.F., Lazarevic D., Lipovich L., Liu J.,  
RA Liuni S., McWilliam S., Madan Babu M., Madera M., Marchionni L.,  
RA Matsuda H., Matsuzawa S., Miki H., Mignone F., Miyake S., Morris K.,  
RA Mortrud-Tabor S., Mulder N., Nakano N., Nakaochi H., Ng P.,  
RA Nilsson R., Nishiguchi S., Nishikawa S., Nori F., Ohara O.,  
RA Okazaki Y., Orlando V., Pang K.C., Pavan W.J., Pavoni G., Pesole G.,  
RA Petrovsky N., Piazza S., Reed J., Reid J.F., Ring B.Z., Ringwald M.,  
RA Roet B., Ruan Y., Salzberg S.L., Sandelin A., Schneider C.,  
RA Schonbach C., Sekiguchi K., Semple C.A., Seno S., Sessa L., Sheng Y.,  
RA Shibata Y., Shimada H., Shimada K., Silva D., Sinclair B.,  
RA Sperling S., Stupka E., Sugtara K., Sultana K., Takenaka Y., Taki K.,  
RA Tammoja K., Tan S.L., Tang S., Taylor M.S., Tegner J., Teichmann S.A.,  
RA Ueda H.R., van Nimwegen B., Vezardo R., Wei C.L., Yagi K.,  
RA Yamashita H., Zabarovsky E., Zhu S., Zimmer A., Hide W., Bult C.,  
RA Grimmond S.M., Tesardale R.D., Liu E.T., Brusic V., Quackenbush J.,  
RA Wahlestedt C., Mattick J.S., Hume D.A., Kai C., Sasaki D., Tomaru Y.,  
RA Fukuda S., Kanamori-Katsuyama M., Suzuki M., Aoki J., Arikawa T.,  
RA Iida Y., Iimura K., Itoh M., Kato T., Kawaji H., Kawagashita N.,  
RA Kawashima T., Kojima M., Kondo S., Kono H., Nakano K., Niimiyi N.,  
RA Nishio T., Okada M., Plessey C., Shibata K., Shikari T., Suzuki S.,  
RA Tasami M., Waki K., Wataniki A., Okamura-Ohno Y., Suzuki H., Kawai J.,  
Hayashizaki Y.;  
RT "The transcriptional landscape of the mammalian genome",  
RL Science 309:1559-1563(2005).  
[3]









**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2006, 17:05:37 ; Search time 199 Seconds  
(without alignments)  
2472.185 Million cell updates/sec

Title: US-09-856-723a-8  
Perfect score: 5489  
Sequence: 1 MKQYASPMPTQTDVTLKFKP.....RTLBNKGKMAKKEKCVLQ 1076

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2589679 seqs, 457216429 residues

Total number of hits satisfying chosen parameters: 2589679

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_8: \*  
1: geneseqp1980s: \*  
2: geneseqp1990s: \*  
3: geneseqp2000s: \*  
4: geneseqp2001s: \*  
5: geneseqp2002s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003bs: \*  
8: geneseqp2004s: \*  
9: geneseqp2005s: \*  
10: geneseqp2006s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1076	AAU74355	AAU74355 Human cyt
2	5486	99.9	1076	ADL13160	ADL13160 Human ste
3	5442.5	99.2	1231	ADX07701	ADX07701 Cyclin-de
4	5266	95.9	1040	AD615640	AD615640 Human str
5	5183	94.4	1581	AAW78788	AAW78788 Human pro
6	5173	94.2	1023	AAW79773	AAW79773 Human pro
7	5173	94.2	1023	AAW79772	AAW79772 Human pro
8	5173	94.2	1023	ABG22883	ABG22883 Novel pol
9	5152	93.9	1023	AAW41108	AAW41108 Human pol
10	5152	93.9	1023	AAW41108	AAW41108 Human pol
11	5152	93.9	1023	AAW41108	AAW41108 Human pol
12	5037	91.8	996	AAW21081	AAW21081 Novel hum
13	5037	91.7	996	AAW21081	AAW21081 Novel hum
14	4817	87.8	960	AAW20269	AAW20269 Novel hum
15	4485	81.7	669	AAW70136	AAW70136 Human hea
16	3151	57.4	669	ABG22882	ABG22882 Novel hum
17	2625	47.8	523	AAW51415	AAW51415 Human wil
18	2405	43.8	487	AAW51417	AAW51417 Human pke
19	2400	43.7	521	AAW78789	AAW78789 Human pro
20	2374	43.3	481	AAW51416	AAW51416 Human pke
21	1045	19.0	221	ABG22880	ABG22880 Novel hum
22	1020.5	18.6	957	ABW70828	ABW70828 Drosophila
23	1009	18.4	202	AAW92539	AAW92539 Human pro

24	989	18.0	479	4	ABG22881	ABG22881 Novel hum
25	945.5	17.2	1523	6	ABW55396	ABW55396 Human MOD
26	924	16.8	1452	7	ADG31335	ADG31335 Human nov
27	897	16.3	175	4	AAW17227	AAW17227 Peptide #
28	897	16.3	175	4	ABW6228	ABW6228 Peptide #
29	897	16.3	175	4	ABW31040	ABW31040 Peptide #
30	897	16.3	175	4	AAW69395	AAW69395 Human bon
31	897	16.3	175	4	AAW57004	AAW57004 Human bra
32	897	16.3	175	4	ABW51067	ABW51067 Human liv
33	897	16.3	175	4	AAW04922	AAW04922 Peptide #
34	897	16.3	175	5	ABG39012	ABG39012 Human pep
35	897	16.3	175	5	ABO60197	ABO60197 Human gen
36	510.5	9.3	996	4	ABW65591	ABW65591 Drosophila
37	507.5	9.2	1010	7	ADG48312	ADG48312 Drosophila
38	498.5	9.1	863	6	ABW98856	ABW98856 Human str
39	498.5	9.1	863	7	ADG31071	ADG31071 Human nov
40	498.5	9.1	863	7	ADG48296	ADG48296 Human MIC
41	498.5	9.1	863	8	ABW82355	ABW82355 Tumour-as
42	493.5	9.0	863	7	ADG48308	ADG48308 Human MIC
43	473.5	8.6	398	4	AAW00864	AAW00864 Human bon
44	473.5	8.6	398	7	ADG33005	ADG33005 Human nov
45	461	8.4	91	4	AAW16170	AAW16170 Peptide #

ALIGNMENTS

RESULT 1	AAU74355	AAU74355 standard; protein; 1076 AA.
ID	AAU74355	
AC	AAU74355	
DT	12-MAR-2002	(first entry)
DE	Human cytoskeleton-associated protein (CYSP) #26.	
XX	Human, cytoskeleton-associated protein; CYSP; autoimmune disorder; cell proliferative disorder; inflammatory disorder; prion disease; vesicle trafficking disorder; gastrointestinal disorder; muscle disorder; neurological disorder; cell motility disorder; reproductive disorder; spinal cord disease; central nervous system disorder; mental disorder; gene therapy; cancer.	
XX	Homo sapiens.	
XX	WO200185942-A2.	
XX	15-NOV-2001.	
PF	03-MAY-2001; 2001WO-US014355.	
XX	05-MAY-2000; 2000US-0201960P.	
PR	08-MAY-2000; 2000US-0202729P.	
PR	05-JUN-2000; 2000US-0209705P.	
PR	07-JUN-2000; 2000US-0210149P.	
PR	21-JUN-2000; 2000US-0213215P.	
XX	(INCY-) INCYTE GENOMICS INC.	
XX	Yue H, Tang YT, Au-Young J, Lu DM, Baughn MR, Hillman JL; Polymers Y, Lal P, Yao MG, Bandman O, Burford N, Batra S, Kearney L; Policy UI.	
PI	WPI; 2002-062248/08.	
DR	N-PSDB; AAS99915.	
XX	New cytoskeleton-associated proteins and polynucleotides, useful for diagnosing, preventing and treating cell proliferative, autoimmune, inflammatory, neurological, cell motility, reproductive and muscle disorders.	
PT	Claim 1; Page 153-155; 194pp; English.	
XX		
PS		

XX The invention relates to human cytoskeleton-associated polypeptides  
 CC (CYSKP) and their associated polynucleotide sequences. The sequences are  
 CC useful in the treatment of disorders associated with overexpression or  
 CC underexpression of CYSKP in a patient. The disorders include cell  
 CC proliferative disorders (such as cancer, actinic keratosis,  
 CC arterioleclerosis, cirrhosis, hepatitis and psoriasis),  
 CC autoimmune/inflammatory disorders (such as, asthma, atherosclerosis,  
 CC osteoporosis, Crohn's disease, rheumatoid arthritis, diabetes mellitus  
 CC and anaemia), vesicle trafficking disorders (such as  
 CC hypercholesterolaemia, diabetes insipidus, Grave's disease and goitre),  
 CC gastrointestinal disorders, prion diseases, neurological disorders (such  
 CC as epilepsy, stroke, cerebral neoplasms, Alzheimer's disease,  
 CC Huntington's disease, Parkinson's disease, amyotrophic lateral sclerosis  
 CC and other motor neuron disorders), cell motility disorders, reproductive  
 CC disorders (such as endometriosis and polycystic ovary syndrome), muscle  
 CC disorders (such as myocardiitis, migraine, hypertension, hypoglycaemia,  
 CC myocardial infarction, epilepsy and muscular dystrophy), spinal cord  
 CC diseases, central nervous system disorders (such as Down syndrome and  
 CC cerebral palsy) and mental disorders (such as anxiety and schizophrenia).  
 CC Sequences AAU74330-AAU74363 represent human CYSKP of the invention.

XX Sequence 1076 AA:

Query Match 99.9%; Score 5486; DB 5; Length 1076;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQYASMPPTQTDVYKLRKPLSKRVVSALQPSLSCIFLRGKATDEDMOSLASVSMQ 60  
 DB 1 MKQYASMPPTQTDVYKLRKPLSKRVVSALQPSLSCIFLRGKATDEDMOSLASVSMQ 60  
 QY ADIGNLDPFEDNEDDENRNOBEKAKITELINKLNFLEAKKDLATVNSNPDPPDA 120  
 DB ADIGNLDPFEDNEDDENRNOBEKAKITELINKLNFLEAKKDLATVNSNPDPPDA 120  
 QY 61 ADIGNLDPFEDNEDDENRNOBEKAKITELINKLNFLEAKKDLATVNSNPDPPDA 120  
 DB 61 ADIGNLDPFEDNEDDENRNOBEKAKITELINKLNFLEAKKDLATVNSNPDPPDA 120  
 QY 121 AELNPFCDPSEBEITETASPRKTEDSFYNNSYNPFKVPQYLYNPFDEBEAVTIKDS 180  
 DB 121 AELNPFCDPSEBEITETASPRKTEDSFYNNSYNPFKVPQYLYNPFDEBEAVTIKDS 180  
 QY 181 PPGSTKRNKIRPVDMSKYLADSKTEBEELDESINPFYEPKSTPPNNLVNPFVLETER 240  
 DB 181 PPGSTKRNKIRPVDMSKYLADSKTEBEELDESINPFYEPKSTPPNNLVNPFVLETER 240  
 QY 241 RVRKAPAPVYLSPKTVLMENTVSAGKDLSTSPSPSPBPVYGRKPNASQSLVWCKE 300  
 DB 241 RVRKAPAPVYLSPKTVLMENTVSAGKDLSTSPSPSPBPVYGRKPNASQSLVWCKE 300  
 QY 301 VTKNVRGVKLTNFTTSWRNGLSFSCAILHFRPDLIDYKSLNPPODIKENKKAADGFAISG 360  
 DB 301 VTKNVRGVKLTNFTTSWRNGLSFSCAILHFRPDLIDYKSLNPPODIKENKKAADGFAISG 360  
 QY 361 ISRLLEPSDMVLAIIPDKLTVMYLYYQIRAHFSGQELNVVQIENSSKSTYKGVNEDT 420  
 DB 361 ISRLLEPSDMVLAIIPDKLTVMYLYYQIRAHFSGQELNVVQIENSSKSTYKGVNEDT 420  
 QY 421 NSGVDOEKFYAELSDLRBEPLQPISGAVDFLSQDDSVFVNDGVESESSEHOTPPDHL 480  
 DB 421 NSGVDOEKFYAELSDLRBEPLQPISGAVDFLSQDDSVFVNDGVESESSEHOTPPDHL 480  
 QY 481 SPSTASPYCARTKSDTEPQKSSQSSGRTSGSDPGICSDNTDSTQAVLLGKRLLKMETL 540  
 DB 481 SPSTASPYCARTKSDTEPQKSSQSSGRTSGSDPGICSDNTDSTQAVLLGKRLLKMETL 540  
 QY 541 ELSDLYVSDKKQWSPFICEETDEQKLTLDIGSNLEKLENSRLSECRSDPESPIKK 600  
 DB 541 ELSDLYVSDKKQWSPFICEETDEQKLTLDIGSNLEKLENSRLSECRSDPESPIKK 600  
 QY 601 TSLSPTSKLGYSTRDLAKKKHSLKOTESDDPDARTLLNHADHSKTVQRLLSROE 660  
 DB 601 TSLSPTSKLGYSTRDLAKKKHSLKOTESDDPDARTLLNHADHSKTVQRLLSROE 660  
 QY 661 ELKERARVLLQOARRDAALAKGNKHTNTATPFCNRQLSDQODERRRQLERARQLIAE 720

DB 661 ELKERARVLLQOARRDAALAKGNKHTNTATPFCNRQLSDQODERRRQLERARQLIAE 720  
 QY 721 ARSGVKNSELPYSGEMAELKERSKASGENDNIEIDTNEEIPGVVGGDELTVLEN 780  
 DB 721 ARSGVKNSELPYSGEMAELKERSKASGENDNIEIDTNEEIPGVVGGDELTVLEN 780  
 QY 781 DLDPEONSKLVLDLKKLLEVPQVNSPSAAOKAVTSSSEODMSGTEDLTERLQK 840  
 DB 781 DLDPEONSKLVLDLKKLLEVPQVNSPSAAOKAVTSSSEODMSGTEDLTERLQK 840  
 QY 841 TTERFNPVVSFSDSTYRKTQLOFSQYINRPEMKORSIOEDTKKNEKKAITEFOR 900  
 DB 841 TTERFNPVVSFSDSTYRKTQLOFSQYINRPEMKORSIOEDTKKNEKKAITEFOR 900  
 QY 901 KPSEDEVLNKGFKDTSGYVGEALALENEQKQIDTRALVBEKRLYIMDGRNTEBBAM 960  
 DB 901 KPSEDEVLNKGFKDTSGYVGEALALENEQKQIDTRALVBEKRLYIMDGRNTEBBAM 960  
 QY 961 MOEWFMVNNKKNALIRMMNQLSLEKEHDEBRYEELNRELAALAEIDMQTEAOKRE 1020  
 DB 961 MOEWFMVNNKKNALIRMMNQLSLEKEHDEBRYEELNRELAALAEIDMQTEAOKRE 1020  
 QY 1021 QLLDELVALVNRKDALVRDLDAQEQABEEDHELTERTLBNKQKAKKEKCVLQ 1076  
 DB 1021 QLLDELVALVNRKDALVRDLDAQEQABEEDHELTERTLBNKQKAKKEKCVLQ 1076

RESULT 2  
 ADLJ3160  
 ID ADLJ3160 standard; protein; 1076 AA.  
 XX  
 AC ADLJ3160;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE Human steroid-induced C3A liver cell protein #134.  
 XX  
 KW Hepatotropic; Gene therapy; Wilson disease; liver disorder;  
 KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US6673549-B1.  
 XX  
 PD 06-JAN-2004.  
 XX  
 PF 12-OCT-2001; 2001US-00976594.  
 XX  
 PR 12-OCT-2000; 2000US-0240409P.  
 XX  
 PA (INCY- ) INCYTE CORP.  
 XX  
 PI Furness LM, Buchbinder JL;  
 XX  
 DR WPI; 2004-068610/07.  
 XX  
 PT Combination useful for preparing a composition for treating liver  
 PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,  
 PT comprises cDNAs that are differentially expressed in response to steroid  
 PT treatment.  
 XX  
 PS Disclosure; SEQ ID NO 889; 141pp; English.  
 XX  
 CC The invention relates to a combination comprising cDNAs that are  
 CC differentially expressed in response to steroid treatment. Also included  
 CC are the following: a high throughput method for using a cDNA to detect  
 CC differential expression of nucleic acids in a sample; and a high  
 CC throughput method of screening molecules or compounds to identify a  
 CC ligand that specifically binds a cDNA. The sample is from a subject with  
 CC Wilson disease and comparison of a standard defines a stage of that  
 CC disease. The high throughput method of screening molecules or compounds  
 CC to identify a ligand that specifically binds a cDNA comprises: combining

CC the combination with molecules or compounds under conditions to allow  
 CC specific binding; and detecting specific binding between each cDNA and at  
 CC least one molecule or compound. The molecules or compounds are regulatory  
 CC proteins. The combination is useful for preparing a composition for  
 CC treating liver disorders associated with steroid therapy, e.g., cirrhosis  
 CC or hepatitis. The present sequence represents a human protein which is  
 CC differentially expressed in steroid-induced C3A liver cells. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification but was obtained in electronic format directly from USPTO  
 CC at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 1076 AA;

Query Match 99.9%; Score 5486; DB 8; Length 1076;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQYASPMPTQTDVTKLKEPLSKRYVSAALQFSLSCIFLRBEGKATDEDMQSLASIVSMKQ 60  
 DB 1 MKQYASPMPTQTDVTKLKEPLSKRYVSAALQFSLSCIFLRBEGKATDEDMQSLASIVSMKQ 60  
 QY 61 ADIGNLDPEDDNEEDENRVQEKAKITELINKLFLDEAKOLATVNSNPPDDDA 120  
 DB 61 ADIGNLDPEDDNEEDENRVQEKAKITELINKLFLDEAKOLATVNSNPPDDDA 120  
 QY 121 AELNPFQDSEBEPITETASPKTEDSFYNNSYNPFKEVQTPQYLNPPDEPAFTIKDS 180  
 DB 121 AELNPFQDSEBEPITETASPKTEDSFYNNSYNPFKEVQTPQYLNPPDEPAFTIKDS 180  
 QY 181 PPGSTKRKNIRPVMSKYLVAADSKTEBEBELDESNPFEYKSTPPNNLVNPFQLETER 240  
 DB 181 PPGSTKRKNIRPVMSKYLVAADSKTEBEBELDESNPFEYKSTPPNNLVNPFQLETER 240  
 QY 241 RYKRAAPAPVLSPTKGLVLENTVSAKGLSTSPKSPSPSPVYLRKKNVAGSGLLVCKE 300  
 DB 241 RYKRAAPAPVLSPTKGLVLENTVSAKGLSTSPKSPSPSPVYLRKKNVAGSGLLVCKE 300  
 QY 301 VTKNRAGKVTNFTTSMRNGLSFCALIHFRPDLIDYKSLNPODIKENNKAYOGFASIG 360  
 DB 301 VTKNRAGKVTNFTTSMRNGLSFCALIHFRPDLIDYKSLNPODIKENNKAYOGFASIG 360  
 QY 361 ISRLLEPSDMVLLAIPDKLTWVLYLQIRAHFSGGELNVQIEENSSTKYKVGNEYETD 420  
 DB 361 ISRLLEPSDMVLLAIPDKLTWVLYLQIRAHFSGGELNVQIEENSSTKYKVGNEYETD 420  
 QY 421 NSVDQEFYALSLDKRBPILQPIGAVDFLSGDDSVFVNDSCVGSSESHQTPDDHL 480  
 DB 421 NSVDQEFYALSLDKRBPILQPIGAVDFLSGDDSVFVNDSCVGSSESHQTPDDHL 480  
 QY 481 SPSRTASPCRRTPKSTPQKQSSGRTSGSDPGICNTSTQVLAQVLLGKKRLKATL 540  
 DB 481 SPSRTASPCRRTPKSTPQKQSSGRTSGSDPGICNTSTQVLAQVLLGKKRLKATL 540  
 QY 541 ELSDLYVADKKKMDSPFICEETDEQKQTLIDIGSNLEKEKLENSRLSECDSPESPICK 600  
 DB 541 ELSDLYVADKKKMDSPFICEETDEQKQTLIDIGSNLEKEKLENSRLSECDSPESPICK 600  
 QY 601 TSLSTFSLGYSYSDLDLAKKKQASLAQTESDPDADRITINHADSSKIVQHRLLSQE 660  
 DB 601 TSLSTFSLGYSYSDLDLAKKKQASLAQTESDPDADRITINHADSSKIVQHRLLSQE 660  
 QY 661 ELKRAARVLQARDAALAKGNKNTNTATPPFCRQSLSDQDERRRQALERRAQLIAE 720  
 DB 661 ELKRAARVLQARDAALAKGNKNTNTATPPFCRQSLSDQDERRRQALERRAQLIAE 720  
 QY 721 ARSGVMSKELPSYSGMAAEKLEKRSKAGDENNDIETNBEIPGFPVGGDELTNLEN 780  
 DB 721 ARSGVMSKELPSYSGMAAEKLEKRSKAGDENNDIETNBEIPGFPVGGDELTNLEN 780  
 QY 781 DLDTPEONSKIVDLKLLKLEVOQVANSPPSSAAQKAVTESSEBDMKSGTEDLRTERLQK 840  
 DB 781 DLDTPEONSKIVDLKLLKLEVOQVANSPPSSAAQKAVTESSEBDMKSGTEDLRTERLQK 840

QY 841 TTERFRNPVFSKSTVARTKLOQSFQYIENRPEMKRORSIOEDTRKNGEKAATTEOR 900  
 DB 841 TTERFRNPVFSKSTVARTKLOQSFQYIENRPEMKRORSIOEDTRKNGEKAATTEOR 900  
 QY 901 KPSEDEVINKGFKTOSQYVVGELALLENBOKOIDTRALVKKRLYLMDTGNTSEBEAM 960  
 DB 901 KPSEDEVINKGFKTOSQYVVGELALLENBOKOIDTRALVKKRLYLMDTGNTSEBEAM 960  
 QY 961 MOEWMFLVNNKQALIRRNQSLLEKEHDLERRYELRLRELPAMALIDEMQTEAKORE 1020  
 DB 961 MOEWMFLVNNKQALIRRNQSLLEKEHDLERRYELRLRELPAMALIDEMQTEAKORE 1020  
 QY 1021 QLLDELVALVNNKDALVRDLDQOEKQAEDEHLERTLBOQNGMAKKEKCVIQ 1076  
 DB 1021 QLLDELVALVNNKDALVRDLDQOEKQAEDEHLERTLBOQNGMAKKEKCVIQ 1076

RESULT 3  
 ADX07701  
 ID ADX07701 standard; protein; 1231 AA.

AC ADX07701;

DT 21-APR-2005 (first entry)

DE Cyclin-dependent kinase modulation biomarker SEQ ID NO 2266.

KW cytostatic; cyclin-dependent kinase; cdk; biomarker.

OS Homo sapiens.

PN MO2005012875-A2.

PD 10-FEB-2005.

PF 29-JUL-2004; 2004WO-US024424.

PR 29-JUL-2003; 2003US-0490890P.

PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

PI Li M, Rudnow BA, Webster KR, Jackson DG, Wong TW;

DR WPI; 2005-163068/17.

DR N-P8DB; ADX07700.

PT Biomarkers useful for predicting or determining the response of a mammal  
 to a cancer treatment comprising administration of a modulator of cyclin-  
 dependent kinase activity.

PS Claim 5; SEQ ID NO 2266; 141bp; English.

This invention describes a novel method of predicting or determining  
 whether a mammal will respond or is responding to an anti-cancer agent  
 that modulates cyclin-dependent kinase (cdk) activity. The method  
 comprises measuring the level of one or more biomarkers selected from  
 2774 biomarkers given in the specification (nucleotide sequence SEQ ID  
 NO.1246 (Genbank E87 W28729) is especially preferred). The method of the  
 invention is utilized in a kit for determining or predicting whether  
 a patient would be susceptible or resistant to treatment by an agent  
 modulating cdk activity. The invention also describes a method for  
 utilizing individualized genetic profiles for treating diseases and  
 disorders based on patient's response and molecular level, specialized  
 microarrays comprising the biomarkers described, antibodies directed  
 against the biomarkers and a cell culture model to identify biomarkers.  
 The cdk modulator is preferably N-5-[[5-(1,1-dimethylethyl)-2-  
 oxazolyl]methylthio]-2-thiazolyl-4-piperidine carboxamide, 0.5-1-  
 tataric acid salt. Note: The sequence data for this patent did not form  
 part of the printed specification, but was obtained in electronic format  
 directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://ftp.wipo.int/pub/published_pct_sequences). This  
 sequence represents a biomarker used in the method of the invention.

XX Sequence 1231 AA;

Query Match 99.2%; Score 5442.5; DB 9; Length 1231;  
 Best Local Similarity 96.6%; Pred. No. 0;  
 Matches 1073; Conservative 1; Mismatches 2; Indels 35; Gaps 1;

QY 1 MKOYASMPPTOTDYKFKPKLSKVVSAAALQPSLSCTFLREGKATDEDMQSLASVSKQ 60  
 DB 121 MKOYASMPPTOTDYKFKPKLSKVVSAAALQPSLSCTFLREGKATDEDMQSLASVSKQ 180  
 QY 61 ADIGNLDPEDNEEDDENRYNOBEKAKIT----- 91  
 DB 181 ADIGNLDPEDNEEDDENRYNOBEKAKITGIYNQNALSSLDEDDODCIKQANMSAK 240  
 QY 92 -----ELINKLANFLDEAEKOLATVNSNPFDDPDPAELNPGDDPDESPITETASPKRTE 145  
 DB 241 SASSEELINKLANFLVEAEKOLATVNSNPFDDPDPAELNPGDDPDESPITETASPKRTE 300  
 QY 146 DSYFNNSYNPPEKEVQTPQYINLPDEPEAFVTIKDSPQSTKRKNI RPYVMSKYLYADSSK 205  
 DB 301 DSYFNNSYNPPEKEVQTPQYINLPDEPEAFVTIKDSPQSTKRKNI RPYVMSKYLYADSSK 360  
 QY 206 TEEBELESNPFYPPKSTPPNNLVNPQLETERRVKRA PAPPVLSPTGVANENTVS 265  
 DB 361 TEEBELESNPFYPPKSTPPNNLVNPQLETERRVKRA PAPPVLSPTGVANENTVS 420  
 QY 266 AGKDLSTSPKPSPIPSVPLGRKPNASOSLWCKEVTKNYGVKITNFTTSMRNGLSFCA 325  
 DB 421 AGKDLSTSPKPSPIPSVPLGRKPNASOSLWCKEVTKNYGVKITNFTTSMRNGLSFCA 480  
 QY 326 ILHHFRPDLIDYKSLNPQDIKENNKAAYDGFASIGISRLBSPDMVLAI PDKLTVMTYL 385  
 DB 481 ILHHFRPDLIDYKSLNPQDIKENNKAAYDGFASIGISRLBSPDMVLAI PDKLTVMTYL 540  
 QY 386 YQIRAHSSGGLANVOI BENSSKSTYKVNGETNTNSSVDDEKYAELSDIKREPELOQP 445  
 DB 541 YQIRAHSSGGLANVOI BENSSKSTYKVNGETNTNSSVDDEKYAELSDIKREPELOQP 600  
 QY 446 ISGAVDELSDODSVFVNDSGVGESESEHQTPDDHLSPTASPYCRTKSDTEPOKSOQS 505  
 DB 601 ISGAVDELSDODSVFVNDSGVGESESEHQTPDDHLSPTASPYCRTKSDTEPOKSOQS 660  
 QY 506 GRTSGSDPGICSNMTDSTQAQVLLGKRLRLKAETLELSDLVYSDKKOMSPPICEETDE 565  
 DB 661 GRTSGSDPGICSNMTDSTQAQVLLGKRLRLKAETLELSDLVYSDKKOMSPPICEETDE 720  
 QY 566 OKQOTLDIGSNLEKKEKLENSRSLERCRSDPESPICKTSLSPYSKLGYSYRDLIAKKKHA 625  
 DB 721 OKQOTLDIGSNLEKKEKLENSRSLERCRSDPESPICKTSLSPYSKLGYSYRDLIAKKKHA 780  
 QY 626 SLRQTESDPADRTTTLNHAHSSKIYHRLLSRQEBLKERARVLLBOARRDAALKAQNKH 685  
 DB 781 SLRQTESDPADRTTTLNHAHSSKIYHRLLSRQEBLKERARVLLBOARRDAALKAQNKH 840  
 QY 686 NNTATPFCNRQJSDQDEERRRQJREARQJLAARSGVMSLEPSYGEMAAKLEKERS 745  
 DB 841 NNTATPFCNRQJSDQDEERRRQJREARQJLAARSGVMSLEPSYGEMAAKLEKERS 900  
 QY 746 KASGDENDNIEIDNEEIPGPFVVGSDDELNLNENDLDPENQNKLYDLKAKLLBPOQ 805  
 DB 901 KASGDENDNIEIDNEEIPGPFVVGSDDELNLNENDLDPENQNKLYDLKAKLLBPOQ 960  
 QY 806 VANSPPSAQAQVATESSEODMKSGTEDLARTERLOKTERFPNPNPVFESKDTVRKTOQSF 865  
 DB 961 VANSPPSAQAQVATESSEODMKSGTEDLARTERLOKTERFPNPNPVFESKDTVRKTOQSF 1020  
 QY 866 SQTENRPEMKRORSIGEDTKKNEKKAITETQRKSEDEVINKGFKDTSQYVVGELAA 925  
 DB 1021 SQTENRPEMKRORSIGEDTKKNEKKAITETQRKSEDEVINKGFKDTSQYVVGELAA 1080  
 QY 926 LENEQOKIDTRAAALVERKRLRYLMDTGRTBESEAMQEWPLVVKKQALLRRNQJLSLE 985  
 DB 1081 LENEQOKIDTRAAALVERKRLRYLMDTGRTBESEAMQEWPLVVKKQALLRRNQJLSLE 1140

QY 986 KEHDLERRRYELNRELAMALIEDMQTEAQKRREQLLDELVALVVKRDALVREDLDAQE 1045  
 DB 1141 KEHDLERRRYELNRELAMALIEDMQTEAQKRREQLLDELVALVVKRDALVREDLDAQE 1200  
 QY 1046 KOAEEDBEHLERTLEONKKGMAKKEKCVIQ 1076  
 DB 1201 KOAEEDBEHLERTLEONKKGMAKKEKCVIQ 1231  
 RESULT 4  
 ADE15640  
 ID ADE15640 standard; protein; 1040 AA.  
 XX  
 AC ADE15640;  
 XX  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human structural and cytoskeleton-associated protein (SCAP) #19.  
 XX  
 KW human; structural and cytoskeleton-associated protein; SCAP;  
 KW arteriosclerosis; atherosclerosis; cirrhosis; hepatitis; myelofibrosis;  
 KW psoriasis; cancer; pneumonia; chronic bronchitis; yellow fever;  
 KW influenza; measles; mumps; HIV; human T lymphotropic virus; rabies;  
 KW gastroenteritis; encephalitis; rubella; epilepsy;  
 KW ischaemic cerebrovascular disease; stroke; cerebral neoplasm;  
 KW Alzheimer's disease; Pick's disease; Huntington's disease; dementia;  
 KW Parkinson's disease; amyotrophic lateral sclerosis; atrophy;  
 KW hereditary ataxia; multiple sclerosis; meningitis; brain abscess;  
 KW prion disease; Creutzfeldt-Jakob disease; Insomnia; neurofibromatosis;  
 KW cerebral palsy; myasthenia gravis; anxiety.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W02003062391-A2.  
 XX  
 PD 31-JUL-2003.  
 XX  
 PF 16-JAN-2003; 2003WO-US001772.  
 XX  
 PR 18-JAN-2002; 2002US-0350702P.  
 PR 25-JAN-2002; 2002US-0351715P.  
 PR 15-FEB-2002; 2002US-0357402P.  
 PR 10-MAY-2002; 2002US-0379880P.  
 PR 17-MAY-2002; 2002US-0381599P.  
 PR 07-JUN-2002; 2002US-0387270P.  
 PR 19-JUL-2002; 2002US-0397125P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 PA  
 XX yue H, Griffin JA, Richardson TW, Tang YT, Thangaveelu K;  
 PI Foreyche IJ, Becha SD, Chavla NK, Hafajia AJA, Swarnakar A;  
 PI Marguis JP, Gotved AE, Baughn MR, Lu DM, Arizizu CS, Kable AD;  
 PI Lee ST, Ramkumar J, Jiang X, Jackson AA, Khare R, Elliott VS;  
 PI Bulloch SA, Xu Y, Lee S, Lehr-Mason PM;  
 XX  
 DR WPI; 2003-671468/63.  
 DR N-PSDB; ADE15671.  
 DR  
 XX New isolated polypeptides useful for treating e.g. cell proliferative  
 PT disorders, viral infections and neurological disorders.  
 XX  
 PT  
 PS Claim 1; SEQ ID NO 19; 357bp; English.  
 PS  
 XX The invention comprises the amino acid and coding sequences of human  
 CC structural and cytoskeleton-associated proteins (SCAP). The SCAP DNA and  
 CC protein sequences of the invention are useful for the diagnosis and  
 CC treatment of: arteriosclerosis, atherosclerosis, cirrhosis, hepatitis,  
 CC myelofibrosis, psoriasis, primary cancer, pneumonia, chronic bronchitis,  
 CC yellow fever, influenza, measles, mumps, HIV, human T lymphotropic virus,  
 CC rabies, gastroenteritis, encephalitis, rubella, epilepsy, ischaemic  
 CC cerebrovascular disease, stroke, cerebral neoplasm, Alzheimer's disease,  
 CC Pick's disease, Huntington's disease, dementia, Parkinson's disease,  
 CC amyotrophic lateral sclerosis, atrophy, hereditary ataxias, multiple



CC scleriosis, meningitis, brain abscess, prion disease, Creutzfeldt-Jakob  
 CC disease, insomnia, neurofibromatosis, cerebral palsy, myasthenia gravis,  
 CC anxiety. The present amino acid sequence represents a human SCAP of the  
 CC invention.

XX Sequence 1040 AA;

Query Match 95.9%; Score 5266; DB 7; Length 1040;  
 Best Local Similarity 96.6%; Pred. No. 2,8e-316;  
 Matches 1039; Conservative 1; Mismatches 0; Indels 30; Gaps 1;

```

QY 1 MKQVAPMPQTQDVAKKFKPLSKVVSALQPSLSCIFLRBQKATDEDMQSLASIVSMKQ 60
DB 1 MKQVAPMPQTQDVAKKFKPLSKVVSALQPSLSCIFLRBQKATDEDMQSLASIVSMKQ 60
QY 61 ADIGLDDPEBNDENRNVQEEKAATITELINKLNLDAEKDLATVNSNPPDDPA 120
DB 61 ADIGLDDPEBNDENRNVQEEKAATITELINKLNLDAEKDLATVNSNPPDDPA 120
QY 121 AALNFGDPDSEEPITETASPKATDSFYNSYNPFKEVQTPQYLNPFDEPAFTIKDS 180
DB 121 AALNFGDPDSEEPITETASPKATDSFYNSYNPFKEVQTPQYLNPFDEPAFTIKDS 180
QY 181 PPQSTKRNIRPVDMSKYLADSSKTEBEELDESMPFEPKSTPPNNLVNPFQELETER 240
DB 181 PPQSTKRNIRPVDMSKYLADSSKTEBEELDESMPFEPKSTPPNNLVNPFQELETER 240
QY 241 RVKRAKAPAPVPLSPKTYLNTVNTVAGKDLSTSPKSPSPSPVLRGRKPAASGLLVWCKE 300
DB 241 RVKRAKAPAPVPLSPKTYLNTVNTVAGKDLSTSPKSPSPSPVLRGRKPAASGLLVWCKE 300
QY 301 VTKNRKGKYNFTTSMRNGLSFCALHFRPDLIDYKSLNPDIKENNKAYDGPASIG 360
DB 301 VTKNRKGKYNFTTSMRNGLSFCALHFRPDLIDYKSLNPDIKENNKAYDGPASIG 360
QY 361 ISRLLEPSDMVLALPDKLVTWTYLYQIRAHPSGQELNVQIEENSRSKTYKVGNEYEDT 420
DB 361 ISRLLEPSDMVLALPDKLVTWTYLYQIRAHPSGQELNVQIEENSRSKTYKVGNEYEDT 420
QY 421 NSSVDQEKFYALSLDKRPELQOPISGAVDFLSDDSVFVNDSGVGSSESHQTPDHL 480
DB 421 NSSVDQEKFYALSLDKRPELQOPISGAVDFLSDDSVFVNDSGVGSSESHQTPDHL 480
QY 481 SPSTASPCRCRTKSTEPQKQSSGRTSGSDPDGICNTSTOATQAVLLGKKRLAKETL 540
DB 481 SPSTASPCRCRTKSTEPQKQSSGRTSGSDPDGICNTSTOATQAVLLGKKRLAKETL 540
QY 541 ELSDLYVSDKKKQMSPPICERTDEQKQLTLDIGSNLEKKEKLENSRSLKESDPSPIKK 600
DB 541 ELSDLYVSDKKKQMSPPICERTDEQKQLTLDIGSNLEKKEKLENSRSLKESDPSPIKK 600
QY 601 TSLSTFSLGYSYSDLDIAKKGHASLRQTESDPADRTTNHADHSSKIYQHRLLSROE 660
DB 601 TSLSTFSLGYSYSDLDIAKKGHASLRQTESDPADRTTNHADHSSKIYQHRLLSROE 660
QY 661 ELKEBARVLLGQARBDALKAQNKNTATPFCRCROISDQODERRRQLEPRAQLAE 720
DB 661 ELKEBARVLLGQARBDALKAQNKNTATPFCRCROISDQODERRRQLEPRAQLAE 720
QY 721 ARSGVQMSLEPSYGEAAEKLKERSKASGDENDNIETDNEIPEGFVGGDELTNLEN 780
DB 721 ARSGVQMSLEPSYGEAAEKLKERSKASGDENDNIETDNEIPEGFVGGDELTNLEN 780
QY 781 DLDPPEQNSKVDLKLKLLLEQVQVANSPPSAAQKAVTESSEQDMKSGTEDLRTERLQK 840
DB 781 DLDPPEQNSKVDLKLKLLLEQVQVANSPPSAAQKAVTESSEQDMKSGTEDLRTERLQK 840
QY 841 TTERPRNVVPSKQSTVARKTQLOSFQYIENRPEMKRORSIOEDPKKNGEKAATTEROR 900
DB 841 TTERPRNVVPSKQSTVARKTQLOSFQYIENRPEMKRORSIOEDPKKNGEKAATTEROR 900
QY 901 KPSBDEVLNKGFKDTQYVVGELALENBQKQIDTRALVVKRLRYLMDTGNTREBEAM 960
DB 901 KPSBDEVLNKGFKDTQYVVGELALENBQKQIDTRALVVKRLRYLMDTGNTREBEAM 960

```

```

DB 865 KPSEDEVLNKGFKDTQYVVGELALENBQKQIDTRALVVKRLRYLMDTGNTREBEAM 924
QY 961 MOEWFMVYKKNQALIRANNOISLIEKEHDLERRVELNRELRAMLAIEDMOKTEAQKRE 1020
DB 925 MOEWFMVYKKNQALIRANNOISLIEKEHDLERRVELNRELRAMLAIEDMOKTEAQKRE 984
QY 1021 QLLDDELVALVYKRDALVRDIDAEQKQAESEDEHLERTLQONKGMKAKKEKCVLQ 1076
DB 985 QLLDDELVALVYKRDALVRDIDAEQKQAESEDEHLERTLQONKGMKAKKEKCVLQ 1040

RESULT 5
AAM78788
ID AAM78788 standard; protein; 1581 AA.
XX
AC AAM78788;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human protein SEQ ID NO 1450.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation.
OS Homo sapiens.
PN WO200157190-A2.
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US004098.
XX
PR 03-FEB-2000; 2000US-00496914.
PR 27-APR-2000; 2000US-00560875.
PR 20-JUN-2000; 2000US-00598075.
PR 19-JUL-2000; 2000US-00620325.
PR 01-SEP-2000; 2000US-00654936.
PR 15-SEP-2000; 2000US-00683561.
PR 20-OCT-2000; 2000US-00693325.
PR 30-NOV-2000; 2000US-00728422.
XX
PA (HISE-) HISEQ INC.
XX
PI Tang YF, Liu C, Dmanac RT, Aeundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao Qa, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
DR WPI; 2001-476283/51.
DR N-PSDB; AAK51921.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.
XX
PS Claim 20; Page 3716-3718; 6221pp; English.
XX
CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication
XX
SQ Sequence 1581 AA;

```

Query Match 94.4%; Score 5183; DB 4; Length 1581;  
Best Local Similarity 67.6%; Pred. No. 6.7e-311;  
Matches 1068; Conservative 1; Mismatches 0; Indels 512; Gaps 1;

QY 8 MPTQDVYKLRKPKLSSKYVSAALQPSLSCIFLRBGKATDEDMQSLASIVSKQADIGMLD 67  
DB 1 MPTQDVYKLRKPKLSSKYVSAALQPSLSCIFLRBGKATDEDMQSLASIVSKQADIGMLD 60  
QY DFEEDNDDDEBNRYNOBEKAAKITELINKNLFLEBAEKDLATVNSNPDDDAELNPF 127  
DB DFEEDNDDDEBNRYNOBEKAAKITELINKNLFLEBAEKDLATVNSNPDDDAELNPF 120  
QY 128 DPDEBEPITETASPRKTEDSPYNNSSYNPFKEVQTPQYLNPFDEBEAFVYTIKOSPQSTKR 187  
DB 121 DPDEBEPITETASPRKTEDSPYNNSSYNPFKEVQTPQYLNPFDEBEAFVYTIKOSPQSTKR 180  
QY 188 KNTRPVMSKYLVDSSKTEBEELDESNPFYEPKSTPPNNLVNPVQELFERRVKKAP 247  
DB 181 KNTRPVMSKYLVDSSKTEBEELDESNPFYEPKSTPPNNLVNPVQELFERRVKKAP 240  
QY 248 APPVLSPRTGVLANENTVSAGKDLSTSPKSPIPSPVLRGKPNASQSLVWCKEYTKNYRG 307  
DB 241 APPVLSPRTGVLANENTVSAGKDLSTSPKSPIPSPVLRGKPNASQSLVWCKEYTKNYRG 300  
QY 308 VKITNFTTSWRNGLSFCALIHFRPDLIDYKSLNPODIKENNKAAYDGFASIGLSRLLEP 367  
DB 301 VKITNFTTSWRNGLSFCALIHFRPDLIDYKSLNPODIKENNKAAYDGFASIGLSRLLEP 360  
QY 368 SDWYLLAIPDKLTMYTYLQIRAHFSGOELNVQIENSSKSTYKVGNYETDINSVDOE 427  
DB 361 SDWYLLAIPDKLTMYTYLQIRAHFSGOELNVQIENSSKSTYKVGNYETDINSVDOE 420  
QY 428 KFYAELSDLRBEPELOPISGAVDFLSQDDSVFVNDSGVGESEHQPDDHLSPTASP 487  
DB 421 KFYAELSDLRBEPELOPISGAVDFLSQDDSVFVNDSGVGESEHQPDDHLSPTASP 480  
QY 488 YCRRTKSDTEPQKSOSSGRTSGSDDPGICSNITDSTOAVLYLGGKRLKATLELSL 547  
DB 481 YCRRTKSDTEPQKSOSSGRTSGSDDPGICSNITDSTOAVLYLGGKRLKATLELSL 540  
QY 548 SDDKKDMSPPPICETEBQKQTLIDIGSNLEKEKLENSRSIECSDPESPICKTSLSPTS 607  
DB 541 SDDKKDMSPPPICETEBQKQTLIDIGSNLEKEKLENSRSIECSDPESPICKTSLSPTS 600  
QY 608 KLGYSYSRDLDLAKGKASLRQTESDPDADRTTLNHADHSSKIYQHRLLSROBELKERAR 667  
DB 601 KLGYSYSRDLDLAKGKASLRQTESDPDADRTTLNHADHSSKIYQHRLLSROBELKERAR 660  
QY 668 VLEQARDAALXAGNKANTNTATPCNRQLSDQDEBERROLERAROLIAEARSQVGM 727  
DB 661 VLEQARDAALXAGNKANTNTATPCNRQLSDQDEBERROLERAROLIAEARSQVGM 720  
QY 728 SELPSYGEAAEKYKERSKASGDENDNIETDNEBIEBGFVGGGDELTYNLENDLTPEQ 787  
DB 721 SELPSYGEAAEKYKERSKASGDENDNIETDNEBIEBGFVGGGDELTYNLENDLTPEQ 780  
QY 788 NSKLVDLKLKKLEVOQVANSPPSSAAOKAVTESSEDMKSGTDLATERLOKTERFRN 847  
DB 781 NSKLVDLKLKKLEVOQVANSPPSSAAOKAVTESSEDMKSGTDLATERLOKTERFRN 840  
QY 848 PUVFSKDSYVRKTOLOFSQYIENRPMKQRSIOEDTKKGNBEKAAITETORPSEDEV 907  
DB 841 PUVFSKDSYVRKTOLOFSQYIENRPMKQRSIOEDTKKGNBEKAAITETORPSEDEV 900  
QY 908 LNKGFQDTSQYVVGELALENEQKQIDTRALVEKRLRYLMDTGRNTEEBEAMQEWPM 967  
DB 901 LNKGFQDTSQYVVGELALENEQKQIDTRALVEKRLRYLMDTGRNTEEBEAMQEWPM 960  
QY 968 VNNKNLIRMNQSLLEKEHDLERRRYELNR----- 999  
DB 961 VNNKNLIRMNQSLLEKEHDLERRRYELNRGISRLLEPSDWYLLAIPDKLTMYTYLQ 1020  
QY 1000 ----- 999

DB 1021 IRAHFGOELNVQIENSSKSTYKVGNYETDINSVDOEKFYAEISDLKREPELOPIS 1080  
QY 1000 ----- 999  
DB 1081 GAVDFLSQDDSVFVNDSGVGESEHQPDDHLSPTASPYCRRTKSDTEPQKSOSSGR 1140  
QY 1000 ----- 999  
DB 1141 TSGSDDPGICSNITDSTOAVLYLGGKRLKATLELSLVDYSDKKDMSPPPICETEBQK 1200  
QY 1000 ----- 999  
DB 1201 LQTLIGSNLEKEKLENSRSIECSDPESPICKTSLSPKLGYSYSRDLDLAKKQASL 1260  
QY 1000 ----- 999  
DB 1261 RQTESDPDADRTTLNHADHSSKIYQHRLLSROBELKERARVLEQARDAALXAGNKANT 1320  
QY 1000 ----- 999  
DB 1321 NTATPCNRQLSDQDEBERROLERAROLIAEARSQVMSSELPSYGEAAEKYKERSKA 1380  
QY 1000 ----- 999  
DB 1381 SGEQNSKLVLDLKLKKLEVOQVANSPPSSAAOKAVTESSEDMKSGTDLATERLOKTE 1440  
QY 1000 ----- 999  
DB 1441 RFRNPVFSKDSYVRKTOLOFSQYIENRPMKQRSIOEDTKKGNBEKAAITETORP 1500  
QY 1000 -----ELRAMLAIEDMOKTEBAOKRREOLLDLVALVNRDALVRDLDQAEKQAEDEHL 1055  
DB 1501 EDEVELRAMLAIEDMOKTEBAOKRREOLLDLVALVNRDALVRDLDQAEKQAEDEHL 1560  
QY 1056 ERTLQNKGMACKKEKCVLQ 1076  
DB 1561 ERTLQNKGMACKKEKCVLQ 1581

RESULT 6  
AAM79773  
ID AAM79773 standard; protein; 1023 AA.  
XX  
AC AAM79773;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human protein SEQ ID NO 3419.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
OS Homo sapiens.  
XX  
PN K0200157190-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001MO-US004098.  
XX  
PR 03-FEB-2000; 2000US-00496914.  
PR 27-APR-2000; 2000US-00560875.  
PR 20-JUN-2000; 2000US-00598075.  
PR 19-JUL-2000; 2000US-00620325.  
PR 01-SEP-2000; 2000US-00654936.  
PR 15-SEP-2000; 2000US-00663561.  
PR 20-OCT-2000; 2000US-00693325.  
PR 30-NOV-2000; 2000US-00728422.  
XX  
PA (HYSR-) HYSEQ INC.



PS Claim 20, Page 328; 6221pp; English.  
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX  
SQ Sequence 1023 AA;  
Query Match 94.2%; Score 5173; DB 4; Length 1023;  
Best Local Similarity 99.7%; Pred. No. 1.6e-310;  
Matches 101; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 63 IGNLDDEEDNEDDENRVNOEKAATBELINKLNFIDEAKDLATNSNPDPDAE 122  
DB :|||||  
DB 10 LANLDDEEDNEDDENRVNOEKAATBELINKLNFIDEAKDLATNSNPDPDAE 69  
QY 123 LNFPGDDSEEPITETASPRKTESFYNNSYNPFKEVOTPOLYLPDPDEBAFVTIKDSP 182  
DB |||||  
DB 70 LNFPGDDSEEPITETASPRKTESFYNNSYNPFKEVOTPOLYLPDPDEBAFVTIKDSP 129  
QY 183 QSTRKKNIRPYDMSKYLVAOSSKTEEBELDESNPFYEKSTPPNNLVNPOLETERRV 242  
DB |||||  
DB 130 QSTRKKNIRPYDMSKYLVAOSSKTEEBELDESNPFYEKSTPPNNLVNPOLETERRV 189  
QY 243 KRKAPAPVLSPTKGVNENTVSAKDLSTPKSPSPISPTLGRKPNASQSLVWCKEVT 302  
DB |||||  
DB 190 KRKAPAPVLSPTKGVNENTVSAKDLSTPKSPSPISPTLGRKPNASQSLVWCKEVT 249  
QY 303 KNYRGVLTNTTTSWRNGLSFCALIHPRPDLIDYKSLNPDDIKENKKAADGFAISGIS 362  
DB |||||  
DB 250 KNYRGVLTNTTTSWRNGLSFCALIHPRPDLIDYKSLNPDDIKENKKAADGFAISGIS 309  
QY 363 RLLEPSDMVLAIAPDKLTVMTYLYQIRAHFSGQELNVQIEBNSKSTYKGVNEYETDINS 422  
DB |||||  
DB 310 RLLEPSDMVLAIAPDKLTVMTYLYQIRAHFSGQELNVQIEBNSKSTYKGVNEYETDINS 369  
QY 423 SVDOEKFYAELSDLKREBELQOPISGAVDFLSQDSYFVNDSGVGESESEKQTPDDHLS 482  
DB |||||  
DB 370 SVDOEKFYAELSDLKREBELQOPISGAVDFLSQDSYFVNDSGVGESESEKQTPDDHLS 429  
QY 483 STASPYCRRTKSDTEPQKSQOSSGRTSGSDPGICSMTDSYQAOVLLGKKRLKAFTLEL 542  
DB |||||  
DB 430 STASPYCRRTKSDTEPQKSQOSSGRTSGSDPGICSMTDSYQAOVLLGKKRLKAFTLEL 489  
QY 543 SDLVSDKKKXDMSPFICEETDEBOKLTLDIGSNLEKELNLSRLBCRSPDESPIKTS 602  
DB |||||  
DB 490 SDLVSDKKKXDMSPFICEETDEBOKLTLDIGSNLEKELNLSRLBCRSPDESPIKTS 549  
QY 603 LSPFSKLGYSYRDLDAKKKGLARQTESPPDARTTLNADHSSKIVQHRLLSRQBEL 662  
DB |||||  
DB 550 LSPFSKLGYSYRDLDAKKKGLARQTESPPDARTTLNADHSSKIVQHRLLSRQBEL 609  
QY 663 KERARVLLBOARBAALKAGKNTNTATPFCNNQLSDQOEERRRDLREBAROLIBAR 722  
DB |||||  
DB 610 KERARVLLBOARBAALKAGKNTNTATPFCNNQLSDQOEERRRDLREBAROLIBAR 669  
QY 723 SGVMSSELPSYGEAAEKLKERSKASGENDNIEIDTNEELPESGFVVGQDELTNLENDL 782  
DB |||||  
DB 670 SGVMSSELPSYGEAAEKLKERSKASGENDNIEIDTNEELPESGFVVGQDELTNLENDL 729  
QY 783 DTEPQNSKLVLDLKLKLLLEVOPQVANSPPSSAAQAVTBESSFQDMKSGTEDLRTERLOKTT 842  
DB |||||  
DB 730 DTEPQNSKLVLDLKLKLLLEVOPQVANSPPSSAAQAVTBESSFQDMKSGTEDLRTERLOKTT 789

QY 843 ERFNPVVFSSKDSYVRKTQLQSFQYIENRPEMRQORSIOEDTKGNEKAATETORKP 902  
DB |||||  
DB 790 ERFNPVVFSSKDSYVRKTQLQSFQYIENRPEMRQORSIOEDTKGNEKAATETORKP 849  
QY 903 SEDEVLNKGFQDSQYVVGELALLENQOKIDTPALVERKRLRYLMDTGRNTEEBEAMQ 962  
DB |||||  
DB 850 SEDEVLNKGFQDSQYVVGELALLENQOKIDTPALVERKRLRYLMDTGRNTEEBEAMQ 909  
QY 963 EWFVLVVKKXALPRRMQLSLBEKHLEBRVYELLNELAMALIEDMQTEAQRRBOL 1022  
DB |||||  
DB 910 EWFVLVVKKXALPRRMQLSLBEKHLEBRVYELLNELAMALIEDMQTEAQRRBOL 969  
QY 1023 LLDLVLVVKKRDALVBDLPAQEKQAEEDHELTERTLEONKGRKAKKEKCVLQ 1076  
DB |||||  
DB 970 LLDLVLVVKKRDALVBDLPAQEKQAEEDHELTERTLEONKGRKAKKEKCVLQ 1023  
RESULT 8  
ABG22883  
ID ABG22883 standard; protein; 1023 AA.  
XX  
XX ABG22883;  
XX  
XX 18-FEB-2002 (first entry)  
XX  
XX Novel human diagnostic protein #22874.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
XX food supplement; medical imaging; diagnostic; genetic disorder.  
XX Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSB-) HYSBQ INC.  
XX  
XX Dmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX N-PSDB; AAS87070.  
XX  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
XX diagnostics, forensic, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX  
XX Claim 20; SEQ ID NO 53242; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensic, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic

CC amino acid sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp://wipo.int/pub/published\_pct\_sequences  
 CC  
 XX Sequence 1023 AA;

Query Match 94.2%; Score 5173; DB 4; Length 1023;  
 Best Local Similarity 99.7%; Pred. No. 1,6e-310;  
 Matches 1011; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

QY 63 IGLDDFEEDNEDDENRVOEKAATLILINKLFLDEAKDLATVNSNPFDDPAE 122
DB 10 LANLDDFEEDNEDDENRVOEKAATLILINKLFLDEAKDLATVNSNPFDDPAE 69
QY 123 LNPFGDPSSEBITETASPRKTEDEFPYNSNYPFEKVOTPOYLNPFDEEAVTIKDSPP 182
DB 70 LNPFGDPSSEBITETASPRKTEDEFPYNSNYPFEKVOTPOYLNPFDEEAVTIKDSPP 129
QY 183 QSTGRKNIRPYDMSKYLVAADSKTEEBELDESNPYEPKSTPPNNLVNPOVLETERRV 242
DB 130 QSTGRKNIRPYDMSKYLVAADSKTEEBELDESNPYEPKSTPPNNLVNPOVLETERRV 189
QY 243 KRKAAPAPVLSPTKGVNMENTVSAGKDLSTSPKSPIPSPVLGRKPNASQSLVWCKEVT 302
DB 190 KRKAAPAPVLSPTKGVNMENTVSAGKDLSTSPKSPIPSPVLGRKPNASQSLVWCKEVT 249
QY 303 KNYGKVTNTNTTSGWRNGLSFCAIILHFRPDLIDYKSLNPODIKENNKAYDGPASIGIS 362
DB 250 KNYGKVTNTNTTSGWRNGLSFCAIILHFRPDLIDYKSLNPODIKENNKAYDGPASIGIS 309
QY 363 RLLEPSDVLVLAIPKLTVMYTYLQIRAHFSGQLNVQIENSSEKSTYKVGNYETDINS 422
DB 310 RLLEPSDVLVLAIPKLTVMYTYLQIRAHFSGQLNVQIENSSEKSTYKVGNYETDINS 369
QY 423 SVDEKFAELSDLKREBELQOPLISGAVDFLSQDDSVVNSGVGESESEHQTPDDHLS 482
DB 370 SVDEKFAELSDLKREBELQOPLISGAVDFLSQDDSVVNSGVGESESEHQTPDDHLS 429
QY 483 STASPYCRRKTSDETPQKSSQSGRTSGSDPGLCSNDSSTQAOVTLGKRLKATETEL 542
DB 430 STASPYCRRKTSDETPQKSSQSGRTSGSDPGLCSNDSSTQAOVTLGKRLKATETEL 489
QY 543 SDLVYSDGKDMSPFICEETDEBQLOFLDIGSNLEKEKLENSRSLBGRSPDESPKIKTS 602
DB 490 SDLVYSDGKDMSPFICEETDEBQLOFLDIGSNLEKEKLENSRSLBGRSPDESPKIKTS 549
QY 603 LSPTSKLGYSYRDLAKKKGASLRQTESPDADRTTLNHAHSSKIVQHRLLSROBEL 662
DB 550 LSPTSKLGYSYRDLAKKKGASLRQTESPDADRTTLNHAHSSKIVQHRLLSROBEL 609
QY 663 KERARVLLEQRDAAALVAKGNKHTNTATPFCNRQLSDQDEERRRLREBARQLIAEAR 722
DB 610 KERARVLLEQRDAAALVAKGNKHTNTATPFCNRQLSDQDEERRRLREBARQLIAEAR 669
QY 723 SGVMSSELPSYGEMAABKLERKSYKSGDENNIRLIDTEKEIPEGVVGGSGELTNLENDL 782
DB 670 SGVMSSELPSYGEMAABKLERKSYKSGDENNIRLIDTEKEIPEGVVGGSGELTNLENDL 729
QY 783 DTPESNKLVDLKLKGLLEVOPOVANSPSAAQKAVTSSSQDMKSGTEDRTERLQKTT 842
DB 730 DTPESNKLVDLKLKGLLEVOPOVANSPSAAQKAVTSSSQDMKSGTEDRTERLQKTT 789
QY 843 ERFNPVVFVSDSTYRKTQLOSPSYIENRPEMKRORSIOEDTKKGNKKAATETQKRP 902
DB 790 ERFNPVVFVSDSTYRKTQLOSPSYIENRPEMKRORSIOEDTKKGNKKAATETQKRP 849
QY 903 SEDVLNKGFDTSQYVVGELAALENKOJDTBAALVERKLRYLMDGRTTERBEAMQ 962
DB 850 SEDVLNKGFDTSQYVVGELAALENKOJDTBAALVERKLRYLMDGRTTERBEAMQ 909
QY 963 EWFMLVNNKALIRRMNQLSLEKEHDLERRRYELLNRELRLAMLAIEDWQTEAQRRRQL 1022
  
```

DB 910 EWFMLVNNKALIRRMNQLSLEKEHDLERRRYELLNRELRLAMLAIEDWQTEAQRRRQL 966  
 QY 1023 LLDLVLVNNKDALVRLDLOEKAEDEDEHLETTLEONKXAKKEKCVLQ 1076  
 DB 970 LLDLVLVNNKDALVRLDLOEKAEDEDEHLETTLEONKXAKKEKCVLQ 1023

RESULT 9  
 AAM41108  
 ID AAM41108 standard; protein; 1023 AA.

AC AAM41108;  
 DT 22-OCT-2001 (first entry)

DE Human polypeptide SEQ ID NO 6039.

KM Human; noctropic; immunosuppressant; cyostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chemokine; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001.

PF 26-DEC-2000; 2000WO-US034263.

PR 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-0048725.  
 PR 25-APR-2000; 2000US-0052317.  
 PR 20-JUN-2000; 2000US-00598042.  
 PR 19-JUL-2000; 2000US-00620312.  
 PR 03-AUG-2000; 2000US-00653450.  
 PR 14-SEP-2000; 2000US-00662191.  
 PR 19-OCT-2000; 2000US-00693036.  
 PR 29-NOV-2000; 2000US-00727344.

PA (HYSB-) HYSBQ INC.

PI Tang YT, Liu C, Auandi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,  
 PI Wang Y, Wang Z, Wehrman T, Xu C, Xue AL, Yang Y, Zhang J, Zhao QA,  
 PI Zhou P, Goodrich R, Drmanac RT;

DR WPI; 2001-442253/47.

DR N-PSDB; AAI60264.

PT Novel nucleic acids and polypeptides, useful for treating disorders such  
 PT as central nervous system injuries.

PS Example 2; SEQ ID NO 6039; 10078bp; English.

CC The invention relates to human nucleic acids (AA157798-AA161369) and the  
 CC encoded polypeptides (AAM8642-AA42213) with noctropic,  
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilization of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification

XX Sequence 1023 AA;

Query Match 93.9%; Score 5152; DB 4; Length 1023;  
 Best Local Similarity 99.3%; Pred. No. 3,1e-309;  
 Matches 1007; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

QY 63 IGLNDLFEEDNEDDENRVNQEKAKITTELLINKLFLDAEKKDLATVNSNPFDDPAAE 122
D 10 LANLDDEEDNEDDENRVNQEKAKITTELLINKLFLDAEKKDLATVNSNPFDDPAAE 69
QY 123 LNPFGDDESEPIETETASPRKTEDSFYNNSSYNPPEKQVTPQYLNPFDPEAFVITIKOSP 182
D 70 LNPFGDDESEPIETETASPRKTEDSFYNNSSYNPPEKQVTPQYLNPFDPEAFVITIKOSP 129
QY 183 QSTRKKRIKRPVDMKSYLYADSSKTEBEELDESNPFEYSKSTPPPNMLVNPQOELETERRV 242
D 130 QSTRKKRIKRPVDMKSYLYADSSKTEBEELDESNPFEYSKSTPPPNMLVNPQOELETERRV 189
QY 243 KRKAAPPVLSPTKGVINENTVSAKDLSTSPKSPPIPSVLGRKPNAOSGLVWCKEVT 302
D 190 KRKAAPPVLSPTKGVINENTVSAKDLSTSPKSPPIPSVLGRKPNAOSGLVWCKEVT 249
QY 303 KNYGVKITNTTTSWRNGLSFCALIHRRPDLIDYKSLNPDIKENNKAADGSAISGIS 362
D 250 KNYGVKITNTTTSWRNGLSFCALIHRRPDLIDYKSLNPDIKENNKAADGSAISGIS 309
QY 363 RLLEPSDMVLTAIPDKLTVMTYLYQIRAHFSGQELNVQIENSSTYKVGNYETDPTS 422
D 310 RLLEPSDMVLTAIPDKLTVMTYLYQIRAHFSGQELNVQIENSSTYKVGNYETDPTS 369
QY 423 SVDEKRYAELSDIKREPELQOPIGAVDFLSQDDVFVNDGVESESEHQTPDDHLS 482
D 370 SVDEKRYAELSDIKREPELQOPIGAVDFLSQDDVFVNDGVESESEHQTPDDHLS 429
QY 483 STASPYGRTKSDTEPOKSOOSSGRTSGSDPGICSNMDSYQAOVLLGKRLLAETTEL 542
D 430 STASPYGRTKSDTEPOKSOOSSGRTSGSDPGICSNMDSYQAOVLLGKRLLAETTEL 489
QY 543 SDLYVSDKKDMSPFICEETDEQKLOTLDIGSNLEKEKLENSSLSECRSPESPICKTS 602
D 490 SDLYVSDKKDMSPFICEETDEQKLOTLDIGSNLEKEKLENSSLSECRSPESPICKTS 549
QY 603 LSPSTKLGYSYRDLDAKKKHAASLRQTESPDADRTTLNHAHSSKIVQHRLLSROBEL 662
D 550 LSPSTKLGYSYRDLDAKKKHAASLRQTESPDADRTTLNHAHSSKIVQHRLLSROBEL 609
QY 663 KEPRVULLEQARRRAALKAGNKHNTNTATPCRNQOLSDQOEERRRLREAROLIABAR 722
D 610 KEPRVULLEQARRRAALKAGNKHNTNTATPCRNQOLSDQOEERRRLREAROLIABAR 669
QY 723 SGVVMSELPSYGEWAAREKLKERSKASGENDNIETDNEELPESGFVVGGBELTNLENDL 782
D 670 SGVVMSELPSYGEWAAREKLKERSKASGENDNIETDNEELPESGFVVGGBELTNLENDL 729
QY 783 DTPEQNSKLVDLKIKLLEVPQVANSPPSSAAQAVTESSBQDKSGTEDLRTERLOKTT 842
D 730 DTPEQNSKLVDLKIKLLEVPQVANSPPSSAAQAVTESSBQDKSGTEDLRTERLOKTT 789
QY 843 ERRRNPVPSKDSVVRKTQOLSFQYIENRPEMRGOSIQGDTCKGNBEKRAITETQRP 902
D 790 ERRRNPVPSKDSVVRKTQOLSFQYIENRPEMRGOSIQGDTCKGNBEKRAITETQRP 849
QY 903 SEDEVLNKGFQDTSQYVVGELAALENBOKQIDTPAAVVEKRLRYLMDTGRTESEBAMQ 962
D 850 SEDEVLNKGFQDTSQYVVGELAALENBOKQIDTPAAVVEKRLRYLMDTGRTESEBAMQ 909
QY 963 EFWPLVANKKALIRRNQQLSLERKHDLERRYELLNLEBLRMLAIEDMQKTEAQKREBL 1022
D 910 EFWPLVANKKALIRRNQQLSLERKHDLERRYELLNLEBLRMLAIEDMQKTEAQKREBL 969
QY 1023 LLEBELVALVKKRDALVNDLDAQEQAAREEDBHLERTLEQNKGNKAKKEKCVLQ 1076
D 970 LLEBELVALVKKRDALVNDLDAQEQAAREEDBHLERTLEQNKGNKAKKEKCVLQ 1023

```

RESULT 10  
 AEA21081  
 ID AEA21081 standard; protein: 1023 AA.  
 XX  
 AC AEA21081;  
 XX  
 DT 11-AUG-2005 (first entry)  
 XX  
 DE Novel human polypeptide SEQ ID NO 1775.  
 XX  
 XX vulnerable; CNS-gen.; gene therapy; diagnostic; forensic; mapping;  
 KW DNA purification; protein purification; osteoarthritis; antiarthritic;  
 KW osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;  
 KW periodontal disease; antiinflammatory; mouth disease; burns; injury;  
 KW peripheral neuropathy; Alzheimer disease; neuroprotective; noctropic;  
 KW degenerative; parkinsons disease; antiparkinsonian; neurological disease;  
 KW cerebrovascular ischemia; cerebroprotective; vasotropic;  
 KW cardiovascular disease; autoimmune disease; immunosuppressive;  
 KW immune disorder; viral infection; virucide; infection; cancer;  
 KW cytostatic; neoplasm.  
 KW  
 OS Homo sapiens.  
 XX  
 PN MO2005049806-A2.  
 XX  
 PD 02-JUN-2005.  
 XX  
 PF 11-MAR-2004; 2004MO-US007412.  
 XX  
 PR 14-MAR-2003; 2003US-00389559.  
 XX  
 PA (NUVE-) NUVELO INC.  
 XX  
 PI Tang TY, Wang J, Wang ZM, Zhang J, Ren F, Zhou P, Ma Y,  
 PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R,  
 PI Wehrman T, Weng G, Boyle B;  
 XX  
 DR WPI; 2005-417730/42.  
 XX  
 XX New polynucleotide encoding a polypeptide with biological activity,  
 PT useful for treating a disease or disorder, e.g. osteoarthritis, burns,  
 PT CNS and peripheral disease, stroke, autoimmune disorders, viral  
 PT infection, or cancer.  
 XX  
 PS Example 3; SEQ ID NO 1775; 500bp; English.  
 XX  
 CC The invention describes a new isolated polynucleotide (1) encoding a  
 CC polypeptide with biological activity comprising: a nucleotide sequence of  
 CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes  
 CC to the sequence of (1) under stringent hybridization conditions; or a  
 CC nucleotide sequence having greater than 98% sequence identity with the  
 CC sequence of (1). Also described are: a(n) (expression)vector comprising  
 CC (1); a host cell genetically engineered to comprise (1) operatively,  
 CC associated with a regulatory sequence that modulates expression of the  
 CC polynucleotide in the host cell; an isolated polypeptide comprising a  
 CC sequence of SEQ ID NOS: 568-1134 (fully defined) where the polypeptide  
 CC is a polypeptide encoded by (1); or a polypeptide encoded by a  
 CC polynucleotide hybridizing under stringent conditions with any one of SEQ  
 CC ID NOS: 1-567; a composition comprising the polypeptide of (3) and a  
 CC carrier; an antibody directed against the polypeptide of (3); a method  
 CC for detecting (1) in a sample; a method for detecting the polypeptide of  
 CC (3) in a sample; a method for identifying a compound that binds to the  
 CC polypeptide of (3); a method of producing the polypeptide of (3); and a  
 CC collection of polynucleotides, where the collection comprising any of at  
 CC least one of SEQ ID NOS: 1-567. (1) is a polynucleotide comprising any of  
 CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological  
 CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:  
 CC 568-1134. All sequences are fully defined in the specification. The  
 CC sequences and methods are useful in diagnostics, forensic, and gene  
 CC mapping, in identifying of mutations responsible for genetic disorders or  
 CC other traits, in assessing biodiversity, and for producing many other  
 CC types of data and products dependent on DNA and amino acid sequences. The





CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological  
 CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:  
 CC 568-1134. All sequences are fully defined in the specification. The  
 CC sequences and methods are useful in diagnostics, forensic, and gene  
 CC mapping, in identifying of mutations responsible for genetic disorders or  
 CC other traits, in assessing biodiversity, and for producing many other  
 CC types of data and products dependent on DNA and amino acid sequences. The  
 CC composition and method are useful for treating a disease or disorder,  
 CC e.g. osteoporosis, osteoarthritis, periodontal disease, burns, CNS and  
 CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,  
 CC autoimmune disorders, viral infection, or cancer. This is the amino acid  
 CC sequence of a novel polypeptide of the invention.

XX  
 XX  
 XX  
 SQ Sequence 1023 AA;

Query Match 93.9%; Score 5152; DB 9; Length 1023;

Best Local Similarity 99.3%; Pred. No. 3.1e-309; Mismatches 4; Gaps 0;

Matches 100%; Conservative 3; Indels 0;

QY 63 IGNLDDEEDNEDDENRNVQEKAKITTELINLFLDEAKDLATVNSNPFDDPAE 122  
 : |||||  
 Db LANLDDEEDNEDDENRNVQEKAKITTELINLFLDEAKDLATVNSNPFDDPAE 69  
 QY 123 LNPFGDSDSEPTETASPRKTEDSFYNSYNPKYVQTPQYLNPFDEPAFTIKDSSP 182  
 : |||||  
 Db LNPFGDSDSEPTETASPRKTEDSFYNSYNPKYVQTPQYLNPFDEPAFTIKDSSP 129  
 QY 183 OSTRKRIKRPVDMKXYLAOSKTEEBELDSENFYVPEKSTPPNNLVNPOLETERRV 242  
 : |||||  
 Db OSTRKRIKRPVDMKXYLAOSKTEEBELDSENFYVPEKSTPPNNLVNPOLETERRV 189  
 QY 243 KRAKAPPVLSPKGVNENTVSAKDLSTSPKSPSPVPLGKRPNASGSLVWCKEVT 302  
 : |||||  
 Db KRAKAPPVLSPKGVNENTVSAKDLSTSPKSPSPVPLGKRPNASGSLVWCKEVT 249  
 QY 190 KRAKAPPVLSPKGVNENTVSAKDLSTSPKSPSPVPLGKRPNASGSLVWCKEVT 249  
 : |||||  
 QY 303 KNYRGVKITNFTTSWRNGLSFCALILHFRPDLIDYKSLMPDIDKENNKAAYDGAISGIS 362  
 : |||||  
 Db KNYRGVKITNFTTSWRNGLSFCALILHFRPDLIDYKSLMPDIDKENNKAAYDGAISGIS 309  
 QY 363 RLLEPSDMVLATPDKLTWMTYLYQIRAHFSGOLNVQIENSSKSTYKGVNETPTNS 422  
 : |||||  
 Db RLLEPSDMVLATPDKLTWMTYLYQIRAHFSGOLNVQIENSSKSTYKGVNETPTNS 369  
 QY 423 SVDQEKYAEISDLKREPELQOPIGAVDFLSDPSFVNDVGSESESEHQTPDDHLSF 482  
 : |||||  
 Db SVDQEKYAEISDLKREPELQOPIGAVDFLSDPSFVNDVGSESESEHQTPDDHLSF 429  
 QY 483 STASPCYCRRTKSDTEPQKQSSGRTSGSDPGICSNVTDSITQAVLLGKRLLAETLEL 542  
 : |||||  
 Db STASPCYCRRTKSDTEPQKQSSGRTSGSDPGICSNVTDSITQAVLLGKRLLAETLEL 489  
 QY 543 SDLYVSPKCKDMSPPTCEETDEKQTLIDIGSLLEKKEKLENSSTLECRSDPESPIKTS 602  
 : |||||  
 Db SDLYVSPKCKDMSPPTCEETDEKQTLIDIGSLLEKKEKLENSSTLECRSDPESPIKTS 549  
 QY 603 LSPSTSKLGYEVRDLADAKKKHASLRTQESDPPDARTTLNHAHSSKIVQHRLLSRQEL 662  
 : |||||  
 Db LSPSTSKLGYEVRDLADAKKKHASLRTQESDPPDARTTLNHAHSSKIVQHRLLSRQEL 609  
 QY 663 KERRAVILLEQARRDALKAGKNTNTATPPCNRQLSDQDEERRRLRERRQLIAEAR 722  
 : |||||  
 Db KERRAVILLEQARRDALKAGKNTNTATPPCNRQLSDQDEERRRLRERRQLIAEAR 669  
 QY 723 SGYKMSLEPSYGEAAAKKERSKASGDENDNTIENNERIPERFVYGGDELINLENDL 782  
 : |||||  
 Db SGYKMSLEPSYGEAAAKKERSKASGDENDNTIENNERIPERFVYGGDELINLENDL 729  
 QY 783 DTEPONSKLVDLKLKLLAVQPOVANSPPSSAAOKAVTESSEBODKSGTEDLRTERLOKTT 842  
 : |||||  
 Db DTEPONSKLVDLKLKLLAVQPOVANSPPSSAAOKAVTESSEBODKSGTEDLRTERLOKTT 789  
 QY 843 ERFRRNPVFSKDSITVRKTOLQSFQYIENRPEMKRORSIOEDTKGNEKKAATTEORKP 902  
 : |||||

Db 790 ERFRRNPVFSKDSITVRKTOLQSFQYIENRPEMKRORSIOEDTKGNEKKAATTEORKP 849  
 QY 903 SEBVLNKGKFDSDQYVVGELALLENQKOIDPRAALVERLRYLMTGTNTEEBEAMQ 962  
 : |||||  
 Db SEBVLNKGKFDSDQYVVGELALLENQKOIDPRAALVERLRYLMTGTNTEEBEAMQ 909  
 QY 963 EWFVLVKNKQALIRNMQLSLEKEHDLERRRYELNRELRLMLAIEDMQTEAKREBOL 1022  
 : |||||  
 Db EWFVLVKNKQALIRNMQLSLEKEHDLERRRYELNRELRLMLAIEDMQTEAKREBOL 969  
 QY 1023 LLDLVALVKNKQALVVDLDAQEKQAEEDHEHLERTLEQNGKMAKKEKCVLQ 1076  
 : |||||  
 Db LLDLVALVKNKQALVVDLDAQEKQAEEDHEHLERTLEQNGKMAKKEKCVLQ 1023

RESULT 12

AEA20268  
 ID AEA20268 standard; protein; 996 AA.

XX  
 AC AEA20268;

XX  
 DT 11-AUG-2005 (first entry)

XX  
 DE Novel human polypeptide SEQ ID NO 962.

XX  
 KW vulnery; CNS-gen.; gene therapy; diagnostic; forensic; mapping;  
 KW DNA purification; protein purification; osteoarthritis; antiarthritic;  
 KW osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;  
 KW periodontal disease; antiinflammatory; mouth disease; burns; injury;  
 KW peripheral neuropathy; Alzheimers disease; neuroprotective; neurologic;  
 KW degeneration; parkinsons disease; antiparkinsonian; neurological disease;  
 KW cerebrovascular ischemia; cerebroprotective; vasodilator;  
 KW cardiovascular disease; autoimmune disease; immunosuppressive;  
 KW immune disorder; viral infection; virucide; infection; cancer;  
 KW cytostatic; neoplasm.

XX  
 OS Homo sapiens.

XX  
 PN W02005049806-A2.

XX  
 PD 02-JUN-2005.

XX  
 PF 11-MAR-2004; 2004MO-US007412.

XX  
 PR 14-MAR-2003; 2003US-00389559.

XX  
 PA (NUVE-) NUVELO INC.

XX  
 PI Tang TY, Wang J, Wang ZW, Zhang J, Ren F, Zhou P, Ma Y;

XX  
 PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R;

XX  
 PI Wehrman T, Meng G, Boyle B;

XX  
 DR WPI, 2005-417730/42.

XX  
 DR N-PSDB; AEA19701.

XX  
 PT New polynucleotide encoding a polypeptide with biological activity,  
 PT useful for treating a disease or disorder, e.g. osteoarthritis, burns,  
 PT CNS and peripheral disease, stroke, autoimmune disorders, viral  
 PT infection, or cancer.

XX  
 PS Claim 20; SEQ ID NO 962; 500bp; English.

XX  
 CC The invention describes a new isolated polynucleotide (I) encoding a  
 CC polypeptide with biological activity comprising: a nucleotide sequence of  
 CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes  
 CC to the sequence of (I) under stringent hybridization conditions; or a  
 CC nucleotide sequence having greater than 99% sequence identity with the  
 CC sequence of (i). Also described are: a(n) (expression)vector comprising  
 CC (i); a host cell genetically engineered to comprise (i) operatively,  
 CC associated with a regulatory sequence that modulates expression of the  
 CC polynucleotide in the host cell; an isolated polypeptide comprising a  
 CC sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide  
 CC is: a polypeptide encoded by (i); or a polypeptide encoded by a





CC encoded polypeptides (AAM38642-AAM42213) with nootropic,  
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC of the invention may be used to treat diseases of the peripheral nervous  
 CC system, such as peripheral nervous injuries, peripheral neuropathy and  
 CC localised neuropathies and central nervous system diseases, such as  
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukemias and  
 CC C.N.S disorders. Note: The sequence data for this patent did not form  
 CC part of the printed specification  
 CC  
 XX

XX Sequence 996 AA;

Query Match 91.7%; Score 5033; DB 4; Length 996;

Best Local Similarity 99.3%; Pred. No. 6.7e-302;  
 Matches 985; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 85 EKAKITELIKLNFLEAEKDLATVNSNPPDDPAALNPGDPDSSEPTTETASPKT 144  
 : :  
 Db 5 KSASSSEBLINKLNFLEAEKDLATVNSNPPDDPAALNPGDPDSSEPTTETASPKT 64  
 QY 145 EDSFYNSYNPFKEVQTPQYLNPFDEPEAFVITIKDSPQSTKRNIRPVMSKTLVADSS 204  
 65 EDSFYNSYNPFKEVQTPQYLNPFDEPEAFVITIKDSPQSTKRNIRPVMSKTLVADSS 124  
 QY 205 KTEBEELDESNPFEYEPKSTPPNNLVNPFQELETERVYKRAKAPVLSPTKGVLENVT 264  
 Db 125 KTEBEELDESNPFEYEPKSTPPNNLVNPFQELETERVYKRAKAPVLSPTKGVLENVT 184  
 QY 265 SAGCDLSTSPRSIIPSPVLGRKNASGSLVWCKEYTKNKGKTKTFTTSNRNGSLFC 324  
 Db 185 SAGCDLSTSPRSIIPSPVLGRKNASGSLVWCKEYTKNKGKTKTFTTSNRNGSLFC 244  
 QY 325 AILHFRPDLIDYKSLNPFQDICKENNKRAYDGFASIGISRLLEPSDMVLAIPDLKLTWY 384  
 Db 245 AILHFRPDLIDYKSLNPFQDICKENNKRAYDGFASIGISRLLEPSDMVLAIPDLKLTWY 304  
 QY 385 LYQIRAHFSGELNVVQIENSRSKTYKVGNVETDTNNSVDQEKRYAELSDLKREPELQ 444  
 Db 305 LYQIRAHFSGELNVVQIENSRSKTYKVGNVETDTNNSVDQEKRYAELSDLKREPELQ 364  
 QY 445 PISGAVDFLSGDDSVFVNDGSGVSESEHQTDPDHLSTSTSPYCRRTKSTTEQKQOS 504  
 Db 365 PISGAVDFLSGDDSVFVNDGSGVSESEHQTDPDHLSTSTSPYCRRTKSTTEQKQOS 424  
 QY 505 SGRSGSDDPGICNTSTOQVLLGKKRLKAETLLESLDYVSDKKKMSPPFICBETD 564  
 Db 425 SGRSGSDDPGICNTSTOQVLLGKKRLKAETLLESLDYVSDKKKMSPPFICBETD 484  
 QY 565 EOKLQTLIDISNLEKEKLENSRSLCERSDPESPICKTSLSPSTKLGYSYSHDLDAKKH 624  
 Db 485 EOKLQTLIDISNLEKEKLENSRSLCERSDPESPICKTSLSPSTKLGYSYSHDLDAKKH 544  
 QY 625 ASLRQTSDDPADDTTLNADHSKTYQHRLLSROBEIKERARVLLQARDDALAKGNK 684  
 Db 545 ASLRQTSDDPADDTTLNADHSKTYQHRLLSROBEIKERARVLLQARDDALAKGNK 604  
 QY 685 HNTVTATPFCNRQSLDQDDEERRRQLRERARQLAARSGVMSBELSPYGMMAEKLEK 744  
 Db 605 HNTVTATPFCNRQSLDQDDEERRRQLRERARQLAARSGVMSBELSPYGMMAEKLEK 664  
 QY 745 SKASGENDNIEIDTNEBIEPFGVVGAGDELTLNENDLDPTEQNSKLVDLKKLLEVP 804  
 Db 665 SKASGENDNIEIDTNEBIEPFGVVGAGDELTLNENDLDPTEQNSKLVDLKKLLEVP 724  
 QY 805 QVANSPSSAAQKAVTESSEQDMKSGTDLRTERLOKTTERRRNVVSKSTVAKTQLOS 864  
 Db 725 QVANSPSSAAQKAVTESSEQDMKSGTDLRTERLOKTTERRRNVVSKSTVAKTQLOS 784

QY 865 PSQYIENRPMKQRSIOEDTKKGNBEKAAITETORPSEDEVLNKGFQTSQYVGE 924  
 Db 785 PSQYIENRPMKQRSIOEDTKKGNBEKAAITETORPSEDEVLNKGFQTSQYVGE 844  
 QY 925 ALENBQKQIDTRALVEKRLRYLMDTGRNTEBEEBAMQEFMLVKNKQALIRRNQSL 984  
 Db 845 ALENBQKQIDTRALVEKRLRYLMDTGRNTEBEEBAMQEFMLVKNKQALIRRNQSL 904  
 QY 985 EKHEDLRRRYELNRLRAMLATIEDQKTEARQRRBQLLDELVALVNXSDALVRDLD 1044  
 Db 905 EKHEDLRRRYELNRLRAMLATIEDQKTEARQRRBQLLDELVALVNXSDALVRDLD 964  
 QY 1045 EKQAEDEHLERTLEBQNKGMKKEKCVLQ 1076  
 Db 965 EKQAEDEHLERTLEBQNKGMKKEKCVLQ 996

RESULT 14  
 AEA20269

ID AEA20269 standard; protein; 960 AA.

XX AEA20269;

DT 11-AUG-2005 (first entry)

XX Novel human polypeptide SEQ ID NO 963.

XX vulnerable; CNS-gen.; gene therapy; diagnostic; forensic; mapping;  
 XX DNA purification; protein purification; osteoarthritis; antiarthritic;  
 XX osteopathic; musculoskeletal disease; osteoporosis; endocrine disease;  
 XX periodontal disease; antiinflammatory; mouth disease; burns; injury;  
 XX peripheral neuropathy; Alzheimers disease; neuroprotective; nootropic;  
 XX degeneration; parkinsons disease; antiparkinsonian; neurological disease;  
 XX cerebrovascular ischemia; cerebroprotective; vasotropic;  
 XX cardiovascular disease; autoimmune disease; immunosuppressive;  
 XX immune disorder; viral infection; viroicide; infection; cancer;  
 XX cytosstatic; neoplasm.

XX Homo sapiens.

XX W02005049806-A2.

PD 02-JUN-2005.

PF 11-MAR-2004; 2004WO-US007412.

XX 14-MAR-2003; 2003US-00389559.

PA (NUVE-) NUVELO INC.

PI Tang TY, Wang J, Wang ZM, Zhang J, Ren F, Zhou P, Ma Y;

PI Ghosh M, Xue A, Asundi V, Zhao Q, Wang D, Goodrich R, Chen R;

PI Wehrman T, Meng G, Boyle B;

DR WPI; 2005-417730/42.

XX N-PsDB; AEA19702.

PT New polynucleotide encoding a polypeptide with biological activity,  
 PT useful for treating a disease or disorder, e.g. osteoarthritis, burns,  
 PT CNS and peripheral disease, stroke, autoimmune disorders, viral  
 PT infection, or cancer.

PS Claim 20; SEQ ID NO 963; 500bp; English.

XX The invention describes a new isolated polynucleotide (I) encoding a  
 CC polypeptide with biological activity comprising: a nucleotide sequence of  
 CC SEQ ID NOS: 1-567 (fully defined); a nucleotide sequence that hybridizes  
 CC to the sequence of (I) under stringent hybridization conditions; or a  
 CC nucleotide sequence having greater than 90% sequence identity with the  
 CC sequence of (I). Also described are: a(n) (expression)vector comprising  
 CC (I); a host cell genetically engineered to comprise (I) operatively,  
 CC associated with a regulatory sequence that modulates expression of the  
 CC polynucleotide in the host cell; an isolated polypeptide comprising a

CC sequence of SEQ ID NOS: 568-1134 (fully defined), where the polypeptide  
 CC is a polypeptide encoded by (1); or a polypeptide encoded by a  
 CC polynucleotide hybridizing under stringent conditions with any one of SEQ  
 CC ID NOS: 1-567; a composition comprising the polypeptide of (3) and a  
 CC carrier; an antibody directed against the polypeptide of (3); a method  
 CC for detecting (1) in a sample; a method for detecting the polypeptide of  
 CC (3) in a sample; a method for identifying a compound that binds to the  
 CC polypeptide of (3); a method for producing the polypeptide of (3); and a  
 CC collection of polynucleotides, where the collection comprising of at  
 CC least one of SEQ ID NOS: 1-567. (1) is a polynucleotide comprising any of  
 CC the sequences of SEQ ID NOS: 1-567 encoding a polypeptide with biological  
 CC activity, which comprises any of the amino acid sequence of SEQ ID NOS:  
 CC 568-1134. All sequences are fully defined in the specification. The  
 CC sequences and methods are useful in diagnostics, forensic, and gene  
 CC mapping, in identifying of mutations responsible for genetic disorders or  
 CC other traits, in assessing biodiversity, and for producing many other  
 CC types of data and products dependent on DNA and amino acid sequences. The  
 CC composition and method are useful for treating a disease or disorder,  
 CC e.g., osteoporosis, osteoarthritis, periodontal disease, burns, CNS and  
 CC peripheral disease, Alzheimer's disease, Parkinson's disease, stroke,  
 CC autoimmune disorders, viral infection, or cancer. This is the amino acid  
 CC sequence of a novel polypeptide of the invention.

XX Sequence 960 AA;

Query Match 87.8%; Score 4817; DB 9; Length 960;

Best Local Similarity 95.8%; Pred. No. 1.4e-288; Matches 950; Conservative 2; Mismatches 4; Indels 36; Gaps 1;

QY 85 EKAAKTELINLNFLEDAEKDLATVNSNPPDDPAALNPPDDSEPIETETASPRKT 144  
 DB 5 KSASSEELINKLNFLEDAEKDLATVNSNPPDDPAALNPPDDSEPIETETASPRKT 64  
 QY 145 EDSFNNSYNPKKQVTPQYLNPPDEPAFAVTKSPPOSTGRKIRIPDMSKTIYADSS 204  
 DB 65 EDSFNNSYNPKKQVTPQYLNPPDEPAFAVTKSPPOSTGRKIRIPDMSKTIYADSS 124  
 QY 205 KTEBELLESNPFYKSTPPNNLVNPPQELERVRVKKAPAPVLSPTKGVNENTV 264  
 DB 125 KTEBELLESNPFYKSTPPNNLVNPPQELERVRVKKAPAPVLSPTKGVNENTV 184  
 QY 265 SAGKDLSTSPKSPSPVYGRKPNASQSLVWCCKEVTNKGKVIKTFETTSWRNGLSFC 324  
 DB 185 SAGKDLSTSPKSPSPVYGRKPNASQSLVWCCKEVTNKGKVIKTFETTSWRNGLSFC 244  
 QY 325 AILHFRPDLIDYKSLNPPDIKENNKKAYDGPASIGISRLLEPSPMTLAIIPDKITVMTY 384  
 DB 245 AILHFRPDLIDYKSLNPPDIKENNKKAYDGPASIGISRLLEPSPMTLAIIPDKITVMTY 304  
 QY 385 LYQIFAHSGOELNVVOIENSSKSTYKVGNETDTNNSVQOEKFAELSDLKREPELQ 444  
 DB 305 LYQIFAHSGOELNVVOIENSSKSTYKVGNETDTNNSVQOEKFAELSDLKREPELQ 364  
 QY 445 PISGAVDFLSQDDSVFVNDGVESESEHQTPDDHLSPTASPCRTKSTDEPOKSQOS 504  
 DB 365 PISGAVDFLSQDDSVFVNDGVESESEHQTPDDHLSPTASPCRTKSTDEPOKSQOS 424  
 QY 505 SGRISGSDPGICSNVTSTQAVLLGKRLKAKFLELSDLVSDKKDKMSPPICEETD 564  
 DB 425 SGRISGSDPGICSNVTSTQAVLLGKRLKAKFLELSDLVSDKKDKMSPPICEETD 484  
 QY 565 EOKIQTLLIGSLTEKELKENSRLTECRSDPSPICKTSLSPSKIGYSRDLDAKKH 624  
 DB 485 EOKIQTLLIGSLTEKELKENSRLTECRSDPSPICKTSLSPSKIGYSRDLDAKKH 544  
 QY 625 ASLRQTESDPDADRFTLLNHAHDSKIVGRHLLSRQBELKERAVLLLEQARRDALAKGNK 684  
 DB 545 ASLRQTESDPDADRFTLLNHAHDSKIVGRHLLSRQBELKERAVLLLEQARRDALAKGNK 604  
 QY 685 HNTNTATPFCNKRQLSDQDEERRQLREBARQLIAEASGVMSGLPSYGEAAAEKLER 744  
 DB 605 HNTNTATPFCNKRQLSDQDEERRQLREBARQLIAEASGVMSGLPSYGEAAAEKLER 664

QY 745 SKASGENDNIEIDTNERIPGPGVVGGBDELTNLENDLDTPEONSKLVDLKLLLEVP 804  
 DB 665 SKASG-----EQNKLVDLKLLLEVP 688  
 QY 805 QVANSPPSAQAQAVTESSEBOMKSGTBDLRTBRLOKTERFRNPPVFSKSTVKTQLOS 864  
 DB 689 QVANSPPSAQAQAVTESSEBOMKSGTBDLRTBRLOKTERFRNPPVFSKSTVKTQLOS 748  
 QY 865 FSOYIENPKMKROSTIEDPKNGEERAAITETORKESEDEVLNKGFEDTSQYVVGELA 924  
 DB 749 FSOYIENPKMKROSTIEDPKNGEERAAITETORKESEDEVLNKGFEDTSQYVVGELA 808  
 QY 925 ALENEKOQIDTPAALVEKRLVYLDYGTNTEEBEAMQEMFVLVKKKVALIRNNQLSL 984  
 DB 809 ALENEKOQIDTPAALVEKRLVYLDYGTNTEEBEAMQEMFVLVKKKVALIRNNQLSL 868  
 QY 985 EKHDLERRYLNLRELPMALIBDMQTEBAQKREQLLDBLVALVKKRDALVDDDAQ 1044  
 DB 869 EKHDLERRYLNLRELPMALIBDMQTEBAQKREQLLDBLVALVKKRDALVDDDAQ 928  
 QY 1045 EKQAEERDEHLERTLEONKGMKKEKCVIQ 1076  
 DB 929 EKQAEERDEHLERTLEONKGMKKEKCVIQ 960

RESULT 15

ADU70136  
 ID ADU70136 standard; protein; 882 AA.

ADU70136;

06-MAY-2004 (first entry)

Human heat mitochondrial protein as a therapeutic target SegID1942.

mitochondrial; human; screening assay; diabetes mellitus;  
 Huntington's disease; osteoarthritis;

Leber's hereditary optic neuropathy; LHON;  
 mitochondrial encephalopathy lactic acidosis and stroke; MELAS;

myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;  
 neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;

osteopathic; ophthalmological; cyostatic.

Homo sapiens.

WO2003087768-A2.

23-OCT-2003.

04-APR-2003; 2003WO-US010870.

12-APR-2002; 2002US-0372843P.

17-JUN-2002; 2002US-038987P.

20-SEP-2002; 2002US-0412418P.

(MITO-) MITOKOR.

(BUCK-) BUCK INST AGE RES.

Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;  
 Warmack DG;

WPI; 2003-845369/78.

Identifying a mitochondrial target for drug screening assays and for  
 treating diseases associated with altered mitochondrial function,  
 PT comprises detecting a modified polypeptide in a sample and correlating  
 with the disease.

Claim 1; SEQ ID NO 1942; 180pp; English.

This invention relates to novel mitochondrial targets that can be used  
 for therapeutic intervention in treating a disease associated with  
 altered mitochondrial function. Specifically, it refers to a method for

CC identifying proteins of the human heart mitochondrial proteome that are  
 CC useful for drug screening assays, as well as therapeutic targets. The  
 CC present invention describes a method for identifying such proteins. The  
 CC can be used in the treatment of various diseases associated with altered  
 CC mitochondrial function including diabetes mellitus, Huntington's disease,  
 CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial  
 CC encephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy  
 CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these  
 CC compositions have neuroprotective, neurotropic, antidiabetic,  
 CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and  
 CC cyostatic activities. This polypeptide sequence is a human heart  
 CC mitochondrial protein of the invention.  
 XX  
 SQ

Sequence 882 AA:

Query Match 81.7%; Score 4485; DB 7; Length 882;

Best Local Similarity 100.0%; Pred. No. 4e-268;

Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 MSKTLVADSSKTEEBELDESNPFPYRKSTPPNNLVNVOELFETRRYKAKAPPVLSLP 254  
 DB 1 MSKTLVADSSKTEEBELDESNPFPYRKSTPPNNLVNVOELFETRRYKAKAPPVLSLP 60  
 QY 255 KTGVLNENTVSAKDLSTSPKSPISPSPVLGRKNASQSLVWCKEVTXNYGKINTFT 314  
 DB 61 KTGVLNENTVSAKDLSTSPKSPISPSPVLGRKNASQSLVWCKEVTXNYGKINTFT 120  
 QY 315 TSMWNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAYDGPASIGISRLLEPSDMVLLA 374  
 DB 121 TSMWNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAYDGPASIGISRLLEPSDMVLLA 180  
 QY 375 IPDLVTWVTYLYOIRAHPSGOELNVVOEENSSKSTYKNGYETDTNSSVDQKPYAELS 434  
 DB 181 IPDLVTWVTYLYOIRAHPSGOELNVVOEENSSKSTYKNGYETDTNSSVDQKPYAELS 240  
 QY 435 DLKKEPELQPISGAVDFLSQDSDVPVNDGSGESESSEHQTPDDHLSPTASPYCRRTKS 494  
 DB 241 DLKKEPELQPISGAVDFLSQDSDVPVNDGSGESESSEHQTPDDHLSPTASPYCRRTKS 300  
 QY 495 DTEPQKQSSGRTSGSDPDICNTDSTQAOVLGKKRLKATLELSDLVYSDKKDM 554  
 DB 301 DTEPQKQSSGRTSGSDPDICNTDSTQAOVLGKKRLKATLELSDLVYSDKKDM 360  
 QY 555 SPPPICETBEQKQOTLDIGSNLEKELENRSLECRSDPESPIKKTSLSPSTKLGYSYS 614  
 DB 361 SPPPICETBEQKQOTLDIGSNLEKELENRSLECRSDPESPIKKTSLSPSTKLGYSYS 420  
 QY 615 RDLDLAKKQASLQTESDPDADRTTLNHADHSSKIYQHRLLSROBELKERAVYLEQAR 674  
 DB 421 RDLDLAKKQASLQTESDPDADRTTLNHADHSSKIYQHRLLSROBELKERAVYLEQAR 480  
 QY 675 RDAALKAQGNNTATPFCRQLSDQODEERRQLRBRARQLTAEARSGVMSLPSYG 734  
 DB 481 RDAALKAQGNNTATPFCRQLSDQODEERRQLRBRARQLTAEARSGVMSLPSYG 540  
 QY 735 EMAAEKLERSKASGDENDNIEIDTNEIEPFGVVGSDDELTNLENDLTPPEONSXLVDL 794  
 DB 541 EMAAEKLERSKASGDENDNIEIDTNEIEPFGVVGSDDELTNLENDLTPPEONSXLVDL 600  
 QY 795 KLKKLLEVOQOVANSPSSAQKAVTESSEDMKSGTEDLRTERLOKTERFRNPFVPSKD 854  
 DB 601 KLKKLLEVOQOVANSPSSAQKAVTESSEDMKSGTEDLRTERLOKTERFRNPFVPSKD 660  
 QY 855 STVAKTQLOSPQYIENRPEKQRQSTIOEDTKKNEKAAITETORXPSDEVLANKGPD 914  
 DB 661 STVAKTQLOSPQYIENRPEKQRQSTIOEDTKKNEKAAITETORXPSDEVLANKGPD 720  
 QY 915 TSOYVVELALAEQEOIDTRALVEKRLRYLMDTGNTBEEBAMQEWPMLVNKKNAL 974  
 DB 721 TSOYVVELALAEQEOIDTRALVEKRLRYLMDTGNTBEEBAMQEWPMLVNKKNAL 780  
 QY 975 IRRNQSLLEKHEHLERRYELLNRELRLAMLAIEDMQTEAQKRQQLLDLVALVNR 1034

DB 781 IRRNQSLLEKHEHLERRYELLNRELRLAMLAIEDMQTEAQKRQQLLDLVALVNR 840  
 QY 1035 DALVRDLDAQEKQAEEDDEHLEERTLEONKGMKKEKCVLQ 1076  
 DB 841 DALVRDLDAQEKQAEEDDEHLEERTLEONKGMKKEKCVLQ 882

Search completed: May 20, 2006, 17:09:12  
 Job time : 207 secs

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 20, 2006, 17:14:58 ; Search time 53 Seconds  
(without alignments)  
1777.037 Million cell updates/sec

Title: US-09-856-723a-8

Perfect score: 5489

Sequence: 1 MKQYASPMPTOTDVKLFKFP.....RTLEQNGKMKAKKEKCVLQ 1076

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 650591 seqs, 87530628 residues

Total number of hits satisfying chosen parameters: 650591

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*

- 1: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/5 COMB.pep:\*
- 2: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/6 COMB.pep:\*
- 3: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/7 COMB.pep:\*
- 4: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/H\_COMB.pep:\*
- 5: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/PTUS\_COMB.pep:\*
- 6: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/RB\_COMB.pep:\*
- 7: /EMC\_Celestra\_SIDS3/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1076	US-09-976-594-889	Sequence 889, App
2	326.5	5.9	3433	US-09-538-092-1136	Sequence 1136, Ap
3	307.5	5.6	2141	US-09-949-016-10918	Sequence 10918, A
4	301.5	5.5	3433	US-09-091-501B-10	Sequence 10, Appl
5	296	5.4	61	US-09-513-999C-7173	Sequence 7173, Ap
6	291.5	5.3	2154	US-08-841-349-4	Sequence 4, Appl
7	291.5	5.3	2154	US-09-431-184A-4	Sequence 4, Appl
8	289.5	5.3	2364	US-09-538-092-1243	Sequence 1243, Ap
9	286	5.2	2388	US-09-695-795A-2	Sequence 2, Appl
10	285	5.2	1026	US-09-949-016-6777	Sequence 6777, Ap
11	285	5.2	1034	US-09-949-016-10870	Sequence 10870, A
12	283.5	5.2	153	US-09-270-767-42273	Sequence 42273, A
13	276.5	5.0	912	US-09-664-958-8	Sequence 8, Appl
14	274	5.0	883	US-09-664-958-10	Sequence 10, Appl
15	270.5	4.9	2008	US-09-091-501B-8	Sequence 8, Appl
16	270.5	4.9	2186	US-09-949-016-10828	Sequence 10828, A
17	270.5	4.9	2349	US-09-538-092-914	Sequence 914, App
18	264	4.8	1078	US-09-248-796A-20284	Sequence 20284, A
19	264	4.8	2662	US-09-595-684B-31	Sequence 31, Appl
20	262.5	4.8	2663	US-09-538-092-1252	Sequence 1252, Ap
21	261.5	4.8	2375	US-09-538-092-1131	Sequence 1131, Ap
22	261	4.8	630	US-09-248-796A-20275	Sequence 20275, A
23	258.5	4.7	1093	US-09-949-016-11535	Sequence 11535, A
24	258.5	4.7	1093	PCT-US93-03077-1	Sequence 1, Appl
25	256	4.7	1976	US-09-538-092-1078	Sequence 1078, Ap
26	255	4.6	10182	US-09-134-001C-3159	Sequence 3159, Ap

27	254.5	4.6	3666	US-09-134-001C-5080	Sequence 5080, Ap
28	254	4.6	2600	US-09-949-016-7309	Sequence 7309, Ap
29	254	4.6	5024	US-09-710-279-2964	Sequence 2964, Ap
30	251	4.6	1588	PCT-US93-07261-11	Sequence 11, Appl
31	251	4.6	1663	PCT-US93-07261-16	Sequence 16, Appl
32	250.5	4.6	1960	US-09-949-016-10872	Sequence 10872, A
33	250	4.6	905	US-09-248-796A-16333	Sequence 16333, A
34	250	4.6	2733	US-09-949-016-11433	Sequence 11433, A
35	250	4.6	3259	US-09-949-016-6507	Sequence 6507, Ap
36	249.5	4.5	1960	US-09-538-092-1077	Sequence 1077, Ap
37	248	4.5	1958	US-10-028-946-4	Sequence 4, Appl
38	248	4.5	2053	US-09-964-956-11	Sequence 11, Appl
39	248	4.5	2054	US-10-028-946-2	Sequence 2, Appl
40	248	4.5	2066	US-09-964-956-9	Sequence 9, Appl
41	247.5	4.5	3878	US-09-914-259-11	Sequence 11, Appl
42	247	4.5	1972	US-08-875-435B-4	Sequence 4, Appl
43	245.5	4.5	1219	US-09-344-624-4	Sequence 4, Appl
44	245	4.5	924	US-09-248-796A-18798	Sequence 18798, A
45	243	4.4	1307	US-09-949-016-7561	Sequence 7561, Ap

#### ALIGNMENTS

```
RESULT 1
US-09-976-594-889
Sequence 889, Application US/09976594
Patent No. 6673549
GENERAL INFORMATION:
APPLICANT: Furness, Michael
APPLICANT: Buchbinder, Jenny
TITLE OF INVENTION: GENES EXPRESSED IN C3A LAYER CELL CULTURES TREATED WITH STEROIDS
FILE REFERENCE: PA-0041 US
CURRENT FILING DATE: 2001-10-12
PRIOR APPLICATION NUMBER: 60/240,409
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 1143
SOFTWARE: PERL Program
SEQ ID NO 889
LENGTH: 1076
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No. 6673549 5202390CD1
US-09-976-594-889

Query Match 99.9% Score 5486; DB 2; Length 1076;
Best Local Similarity 99.9% Pred. No. 0;
Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKQYASPMPTOTDVKLFKFP...SALQFSLSCIFLRGKATDEDMQSLASLVSMKQ 60
|||||
DB 1 MKQYASPMPTOTDVKLFKFP...SKKVVSAALQFSLSCIFLRGKATDEDMQSLASLVSMKQ 60
|||||

QY 61 ADIGNLDFEEDNEDDENRYNOSBKAATTELINKLNFLEABEKDLATVNSNPPDDPA 120
|||||
DB 61 ADIGNLDFEEDNEDDENRYNOSBKAATTELINKLNFLEABEKDLATVNSNPPDDPA 120
|||||

QY 121 AELNPFQDDEEPTTETASPRKTEDSFYNSNYPKFEVOTPOYLNPDEBEAFVITIDS 180
|||||
DB 121 AELNPFQDDEEPTTETASPRKTEDSFYNSNYPKFEVOTPOYLNPDEBEAFVITIDS 180
|||||

QY 181 PPOSTKRKNIRPVMSKTYADSSKTEBEBELDBSNPFYEPKSTPPNNLVNPFVLELTER 240
|||||
DB 181 PPOSTKRKNIRPVMSKTYADSSKTEBEBELDBSNPFYEPKSTPPNNLVNPFVLELTER 240
|||||

QY 241 RYKKAAPAPPLSKTYGLMENTYSAGKDLSTSRKSPRISPPVGRKRNASQSLVWKE 300
|||||
DB 241 RYKKAAPAPPLSKTYGLMENTYSAGKDLSTSRKSPRISPPVGRKRNASQSLVWKE 300
|||||

QY 301 VTKYRGVKTNTFTSWRNGLSFCALLHHFRPDLIDYKSLNPQDIKENNKAYGDFASIG 360
|||||
```

```
Db 301 VTKAVRGVLTNFTTSMRNG..SFCALIHHPADLIDYKSLNPQIKENKKAAYOGFASIG 360
Qy 361 ISRLLESDWVLAIIPKLTVMYLYQIRAHFSGQELNVVQIENSSSTYKNGNYEDT 420
Db 361 ISRLLESDWVLAIIPDLVTMYLYQIRAHFSGQELNVVQIENSSSTYKNGNYEDT 420
Qy 421 NSSVDQEKFAELSDLKREPELQPISGAVDFLSQDVSFVNDGSGVSESEHQTDPDDL 480
Db 421 NSSVDQEKFAELSDLKREPELQPISGAVDFLSQDVSFVNDGSGVSESEHQTDPDDL 480
Qy 481 SPSTASPYCRRTKSDTEPQKSGSSGRTSGSDDPGICSNUTSTQAVLIGKRLIKAE 540
Db 481 SPSTASPYCRRTKSDTEPQKSGSSGRTSGSDDPGICSNUTSTQAVLIGKRLIKAE 540
Qy 541 ELSPDLYSDKKKMSPPICERTBQKQLTLDISNLEKELENSRSLSECSDESPYK 600
Db 541 ELSPDLYSDKKKMSPPICERTBQKQLTLDISNLEKELENSRSLSECSDESPYK 600
Qy 601 TSLSPSTKLGYSYGRDLDAKKKHAASLRQTESDPDADRTTLNHADHSKIVQHRLLSRO 660
Db 601 TSLSPSTKLGYSYGRDLDAKKKHAASLRQTESDPDADRTTLNHADHSKIVQHRLLSRO 660
Qy 661 ELKERAVLLEQARDAALKAQKNTNTATPCNRLSDQODEERRRQLEBARQILAE 720
Db 661 ELKERAVLLEQARDAALKAQKNTNTATPCNRLSDQODEERRRQLEBARQILAE 720
Qy 721 ARSGVKNSELPSTYGEAAEKLEKSKASGDENDNIEIDTNEIEIEGPGVGGDELTYLEN 780
Db 721 ARSGVKNSELPSTYGEAAEKLEKSKASGDENDNIEIDTNEIEIEGPGVGGDELTYLEN 780
Qy 781 DLDPPEONSKLVLDLKKLEVPQVANSPPSSAQKATVSSSEODMKSCTEDLTERLOK 840
Db 781 DLDPPEONSKLVLDLKKLEVPQVANSPPSSAQKATVSSSEODMKSCTEDLTERLOK 840
Qy 841 TTERFRNPVVSFQSTYRKTQLQSFQYIENRPMKQORSIQEDTKKGNBEKAAITETOR 900
Db 841 TTERFRNPVVSFQSTYRKTQLQSFQYIENRPMKQORSIQEDTKKGNBEKAAITETOR 900
Qy 901 KPSPDEVLNKGFKDTQSYVGEELALENBOKQIDTRALVYKRLRYLMDTGRNTEEBEAM 960
Db 901 KPSPDEVLNKGFKDTQSYVGEELALENBOKQIDTRALVYKRLRYLMDTGRNTEEBEAM 960
Qy 961 MOEWFMLVNNKGNALIRRMNQLSLKEHDEHRYELNRELRLAMLAIEDMQTEAQRRE 1020
Db 961 MOEWFMLVNNKGNALIRRMNQLSLKEHDEHRYELNRELRLAMLAIEDMQTEAQRRE 1020
Qy 1021 QLLDELVALVNNKDALVRDLDAQEKQAEEDDEHLERTLEQNGKMAKKEKCVLQ 1076
Db 1021 QLLDELVALVNNKDALVRDLDAQEKQAEEDDEHLERTLEQNGKMAKKEKCVLQ 1076
```

```
RESULT 2
US-09-538-092-1136
; Sequence 1136, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Manesfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538, 092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuratasegFormatter Version 0.9
; SEQ ID NO 1136
; LENGTH: 3433
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (0)..(0)
; OTHER INFORMATION: Polypeptide Accession Number P46939
US-09-538-092-1136
Query Match 5.9%; Score 326.5; DB 2; Length 3433;
Best Local Similarity 19.8%; Pred. No. 2.6e-12;
Matches 197; Conservative 159; Mismatches 336; Indels 303; Gaps 37;
;
Qy 286 RKNVASQSLVWCSEVTKNRYGKVTNFTTSMRNGSFCALIHHPADLIDYKSLNPQDI 345
Db 148 QQTNSEKILLISWROQTRPYQVAVNLFTTSMWDLAFNVLNHRKEDLSWDCVWVWSP 207
Qy 346 KENKKAAYOGFAS-IGISRLLESDWVLAIIPDLVTMYLYQI- 388
Db 208 IERLEHAFSAQTYLIGIKLLDPED-VAVRLPDKKSITMYLTSIFEVLPOQVITDAIREV 266
Qy 389 ----RAHFGQELNVVQIENSSSTYKNGNYEDT--TNSVDQEKFAELSDLKREPR 441
Db 267 ETLPRKXKCEBEALINIGSTAPDEEHSPRAETPSTVTEVMDLDSYQIALREV----- 321
Qy 442 LQOPISGAVDFLSQDVSFVNDGSGVSESEHQTDPDDLSPSTAS- 488
Db 322 LTMLSAEDTFOEQDD--ISDD-VBEVKKQFATHEAFMMLTANHOSSVGSVLQAGNQLI 377
Qy 489 CRRTKSDTEPQKQ-----QSSGRTSGSDDPGICSNUTSTQAVL----- 528
Db 378 TQGLTSDSEEFIEQWNTLNAEALRVESMDRQSLHD-VLMELQKKQLQSLAMLT 435
Qy 529 LGKRLIKAEITLSD-----LYVSDKKKDMSP 556
Db 436 LTSEKIQKMETCLLDLVKSLQKLEBHKSLQSDLEABQKANSITMNVIVDENGESA 495
Qy 557 PFI-----CESTDQ--KLTLDIGSNLEKEKLENSRSLSECSDESPYK 599
Db 496 TAILLEDQLKIGRWTAVCWTEERWNRLOEINI--LMQELLEE--OCLLKAWLTEX 548
Qy 600 KTSLSPTSKLGYSYGRDLDAKKKHAASLRQTESDPDADRTTLNHADHSKIVQ----- 652
Db 549 BEALNKVQTSNFDQKELSVVRLLAIKE--DMEMKQTL--DLSIEIGQDVGQLD 602
Qy 653 ----HRLSROBELKER-----ARVLLQAR-RDAA 678
Db 603 NSKASKKINSDESELQRMWSLVQRLEDSNGYQAVAKLGMSQIPQKDLLETFRVAREQA 662
Qy 679 LKAGKNTNTATPCNRLSDQODEERRRQLEBARQIL--IAEASGVNMSLPSTYGE 735
Db 663 ITTKSGQELPPPPPKKQI--HVDIEAKKFPALISABLNTLTKMTALQTTEIKYMK 720
Qy 736 MA-ABKLEKSKASGDENDNIEIDTNEIEIEGPGVGGDELTYLENDLDRPEONSKLVLD 794
Db 721 MDPTSEKKQLKA-----LEKQREHRLPR-----ADELNTQGLIVQMGKEGLPTE 767
Qy 795 KLKCLEVPQVANSPPSSAQKATVSSSEODMKSCTEDLTERLOKTTERRNPVVSFQ 854
Db 768 EIKGVLE-----KVSSEMKNVSGHJEDLEK-----IQIQED 799
Qy 855 STYKQQLQSFQYIENRPMKQORSIQEDTK-----KGNBEK 892
Db 800 INAVFQDLBEVYIKTKEMVWHTSISBSRQSLPSLKDSCORELTNLGLHPKIEMAR 859
Qy 893 AATTEQKSEDEVLNKG-----KDTQYVVEELAALENBOKQIDTRAA 938
Db 860 ASCSALMSQSPADPFQGRGDFLGRYQAVQAEVDRQDH-----LEVELGQPGHAY 912
Qy 939 LVEKRLRYLMDTGRNTE-----EEEMNQEWFMVLVNNKGNALIRRM 978
Db 913 L--ETLKTLDVANDENKQVSLNVLNDLAKYKALQEKTKTDE--ILENQFALHKL 968
Qy 979 NQSLLEK--EHLDEHRYELNRELRLAMLAIEDMQTEAQRREQLLDLVL----- 1028
Db 969 EETKALEKNVHPRVEKLVQGFDDVQG-----KMNKLVVSKDLLEIEIALTLBAFEA 1023
```

Qy 1029 --ALVNRKDALVRLDAQEKQAEEDHEHLETLQ 1061  
Db 1024 DSTVLEKMDGVKDFLPMKQAAQSPDAGLQRLDQ 1058

RESULT 3  
US-09-949-016-10918  
; Sequence 10918, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10918  
; LENGTH: 2141  
; TYPE: PRT  
; ORGANISM: Human  
US-09-949-016-10918

Query Match 5.6%; Score 307.5; DB 2; Length 2141;  
Best Local Similarity 20.0%; Pred. No. 2,4e-11;  
Matches 196; Conservative 165; Mismatches 318; Indels 303; Gaps 45;

Qy 285 GRKP-NASOSLLVWCKEYTKYRGVKTINFTTSMNGLSFCALIHFRPDLIDYKSLNPQ 343  
Db 173 GRETSADALLWCMQMTAGPHNVNTFTSSWKGALFNLHKKRPDLIDPKLXDS 232

Qy 344 DIKENNKAYD-GFASIGISRLLESDVNLAIPIKLTVMYTYLYDIRAHFSGQELNVQI 402  
Db 223 NARHNLBHAFFVAERQGLIIPLDPEP-VFTENPDEKIIITVVAFYHYFS--KKKVLAV 289

Qy 403 EENSSTKYKGVNYETDINSSVDOKFYAIBSDL----- 436  
Db 290 E--GKRVGVYIDHAIETEKMI--EKYSGLASDLTWTIEQITVLNSKKFANSLTGVQOQ 344

Qy 437 -----KRPELQOPISGAVDFLSQDDSVFVNDGSGVSESEHOTPPDHLSPSTAS 486  
Db 345 LQASTVATVEKPRFQF--KGNLEVL-----LFTIQSRMANNOKVTPHD--GKLV 394

Qy 487 PYCRTKSDTEPQSQSSGRTSGSDPDGICSTNDSTOAQVILGKRLIKATYLE-LSDL 545  
Db 395 DINRAMESLERAEYRR-----LALNELIRQEKLEQLARR 430

Qy 546 YVSDKKKMSPPFCE-----ETD-----RQKQTL-DI 573  
Db 431 F--DKAAMRETWSENOPLVAQDNFGYDLAANBAKKGHAIETDTAAVERVAALEDL 488

Qy 574 GSNLEKEKLNSRSLSECSDESPKIKTSLSPSTKLSGYSRDLDLAKKHAASLQTESD 633  
Db 489 AQELEKEKHYHQKRTAKKDNILRL-----WSYIQ--ELQSRQRETTIAL 534

Qy 634 PDADRTTLNADHASKIYQHRU-----LSRQESLKRARAVLLEQARRDALAKGNKN 686  
Db 535 QKLPQDMLHSIDWDEIRVAHLISAFGKHLLEVEDLQKHKLW-----EADTAIQ-GDKVK 589

Qy 687 TMTATPF-----CNROL-----SPQODE-----RRQLEBAROL----- 717  
Db 590 AITTAATLKFTGSKYQPCDPQVITDRISHLEQCEBELSNMAAGRAQL--EOSKELMKFTW 648

Qy 718 -IAEARSGVKNSELPSYGEAAEK-----LKERSKAGDENNDIEIDTNEBRIPE--GF 767

Db 649 EMDAEASWIKKE-QIYSSLDYKDLTSVLLIQKHKAFEDBLGDLNLEQIFQEAHGM 707  
Qy 768 VVGGEDELTNLENLDP--RQNSKLYDLKXKLELVQPVANSESSAAQKAVTESSQD 825  
Db 708 VA-----RKQGFHPOIARIKESVQMDQLDLAFCCKNLQDENPFQFGDADD 758

Qy 826 MMSGTEDLRTERLQKTERFPNVPVFSKDSYVRKTQLOSFGQYLENREMKRORSIQBDT 885  
Db 759 LKAWLQD-----AHRLLSGEDVQDBGATBALGKHNDPFL--ELBESRGVMEHL 806

Qy 886 KKGNEBKALITETQKQSEDEVLNKGFKDYSQYVGEALALENROKQIDTPAALVEKRL 945  
Db 807 EQ-----QAQGFPEE-----FRD-SPDVTNRLQALRELYQVVAQADLRQRLQ 849

Qy 946 YLMDGRTSESEAMQEWFLVYKKAALIRRMQSLSEKGNLERYELLNRELRML 1005  
Db 850 EALDLYTVFGSTDA-CELM--WGEKEXKLAEMEPDTELBLEVVQHFDTLDDQEMKTL 905

Qy 1006 ALED-----WQ--KTEAQKREQL----- 1022  
Db 906 TQIDGVNLAAANSVYESGHPREVRKYQODHLNTRMQAFQTLVSRREAVDSALVHNVCV 965

Qy 1023 -----LQDE-----LVALVNRDALVVD-----LDAQEKQAE- 1050  
Db 966 DEESTSKWITDKTKVVESTKDLGRDLAIAIQKRLSGLERDVAIQARVDALERESQOL 1025

Qy 1051 EDEHLERTLBNQKMAKKEK 1072  
Db 1026 MDSHEQ--KRDIGQRQKHLE 1045

RESULT 4  
US-09-091-501B-10  
; Sequence 10, Application US/09091501B  
; Patent No. 6518413  
; GENERAL INFORMATION:  
; APPLICANT: Tinsley, Jonathan M  
; APPLICANT: Davies, Kay E  
; TITLE OF INVENTION: Utrrophin gene expression  
; FILE REFERENCE: 620-42  
; CURRENT APPLICATION NUMBER: US/09/091,501B  
; CURRENT FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: PCT/GB96/03156  
; PRIOR FILING DATE: 1996-12-19  
; PRIOR APPLICATION NUMBER: GB 9525962.8  
; PRIOR FILING DATE: 1995-12-19  
; PRIOR APPLICATION NUMBER: GB 9615797.9  
; PRIOR FILING DATE: 1996-07-26  
; PRIOR APPLICATION NUMBER: GB 9622174.2  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 3433  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (239) .. (250)  
; OTHER INFORMATION: Description of Artificial Sequence: Full length  
; OTHER INFORMATION: utrophin construct; Xaa = unknown  
US-09-091-501B-10

Query Match 5.5%; Score 301.5; DB 2; Length 3433;  
Best Local Similarity 19.6%; Pred. No. 1.1e-10;  
Matches 190; Conservative 165; Mismatches 363; Indels 251; Gaps 36;

Qy 286 RKPASOSLLVWCKEYTKYRGVKTINFTTSMNGLSFCALIHFRPDLIDYKSLNPQDI 345  
Db 148 QQNTSKILSVARQTPRYSQVAVNLFTTSMNGLAFNVLHKKRPDLIDYKSLNPQDI 207

Qy 346 KENNKAYD-GFASIGISRLLESDVNLAIPIKLTVMYTYLYDIRAHFSGQELNVQI 402  
Db 346 KENNKAYD-GFASIGISRLLESDVNLAIPIKLTVMYTYLYDIRAHFSGQELNVQI 402

```

Db      208 IERLEHAFSKAHYIAGIEKLDLPEDVAHVLPEXXXXXXXVYVLEFQVTDIDAREVE 267
Qy      383 TYLQIDRAHFGQELNV---VQIEENSSKSTYKVGNYETDNTSSVDQSKFAEISDLKRE 439
Db      268 TLPRKYKKECEEEBHHIQAVALABEGSP-----RAEPSTVTEVDMDLDSYQIALE 319
Qy      440 PELQOPIGAVDFLSQDDSVFVNDGVESESEHQTPDDHLSPTASPCRTKSDTEPQ 499
Db      320 EVLFWLMSAEDTFOEQDD---ISDD-VREVEQGFATHTFTFWMLJA-----H 362
Qy      500 KSOQSSGRTSGSD--DEIGCNTSTQAO---VILGKRLIKKATLE---LSDLVSD 549
Db      363 QSSVSVYLQAGNQMLTQGTLSSEEFIEQEQMTLINAWEALRVESMERQRLDAMEL 422
Qy      550 KKK---DMSPPFICEETDEQKLOTLDIGSNLE--KEKLNGRSLECSDSPSFKTSLIS 604
Db      423 QKKLOQLSSWLTATBERIQMESPPGDDLPISQKLIQEKHSIQNDLEAQ-VKVNSLT 481
Qy      605 -----PTSKLGYSSRDLDAKKKASLQ----- 629
Db      482 HMVYIVDENGESATVALLDQQLGGERMTAVCMTETERNMRLQEISILMQELLEQCL 541
Qy      630 ----TESDPADRT-TIANHADHSSKIVQHRLSROEIKERARVLEQARDALAKGAK 684
Db      542 EAMLTKEEBALNKYQTSNFKDQKEISVVRRLALIKEDMCKRQTLDLS-----EIGOD 596
Qy      685 HNTVATPFCNRQLSDQDDEERRR-----QLRRARQULIARSGVMSELPSYGEMAAE 739
Db      597 VGQLSPKSKSKKMSSEELTORWDSLQRLSDSNQVTOAVAKLQSGQIPQKDLSTV 656
Qy      740 KLERKAS-----GDENNIIEIDTNEBIEPEGFVVGGD-----E 774
Db      657 HVRKKGAVKPKQELPPPLGPKRQIHVDI--EAKKFFDAISALMLMLKMTAIGTTE 714
Qy      775 LTNLENDLTPBQNSKLVDELKLLKLEVOQVANSFSSAAQKAVTESSEODMKGSTEDLR 834
Db      715 IKEXYMKQDTSEMKKKIKALE-KEQREIRP-ADELAQTCQILVEQ-----MGKGLP 765
Qy      835 TER-----LOKTERFRN-----PVVFSKQSTYRKQLOLQSFQYIENRPEMKQRS 880
Db      766 TERIKNVLEKVSSEWKVNSQHLBDELRKIQLOEINAVFYQDLDELEKVIKTEEMVWHTS 825
Qy      881 IOEDTK-----KGNREKAITETORKEPSEDEVLNKGF----- 912
Db      826 ISSBSRQSLPSLKQSCQBELTNLGLHPKITEMARASQALMSQSDADPVGROGDSFLGR 885
Qy      913 -----KTSQYVVBELALENQKQIDTRALVETRLRYLMDTGRNTE----- 955
Db      886 YQAVQEAVEDROQH-----LENELKGQPGHAYL--ETLKTLDVLANDSENKAQVSLNV 936
Qy      956 -----BEEAMWQEWFMVLVKKQNALIRRNQOLSLLEK--EHLERRYBELANBLR 1002
Db      937 LNDLAKYKELKQEKKTIDE--ILENQKPALHKLAEETKALEKVNHPVEKLYKQEPDVO 994
Qy      1003 AMLAIEDMQTEAQKREQLLDLDELV-----ALVNRDALVLDLAQEKQAEED 1052
Db      995 G-----KMNGLKVLVSDFHLHLEBIALTLRAFEADSVVIEKMDGVDFLMKQQAAGDD 1049
Qy      1053 EHLERTLEQ 1061
Db      1050 AGLQROLQ 1058

```

```

RESULT 5
US-09-513-999C-7173
; Sequence 7173, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961

```

```

; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513, 999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 7173
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 16
; OTHER INFORMATION: Xaa=Ile or Lys or Arg or Thr
; FEATURE:
; NAME/KEY: UNSTRE
; LOCATION: 17
; OTHER INFORMATION: Xaa=Asp or Glu
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 39
; OTHER INFORMATION: Xaa=Asp or Glu
; US-09-513-999C-7173

Query Match          5.4%; Score 296; DB 2; Length 61;
Best Local Similarity 95.1%; Pred. No. 1,1e-12;
Matches 58; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      727 MSELPSYGEAAEKLEKRSKASGDENDNIEIDTNEBIEPEGFVVGGDDELTNLENDLTPR 786
Db      1 MSELPSYGEAAEKLEKRSKASGDENDNIEIDTNEBIEPEGFVVGGDDELTNLENDLTPR 60
Qy      787 Q 787
Db      61 Q 61

RESULT 6
US-08-841-349-4
; Sequence 4, Application US/08841349B
; Patent No. 595594
; GENERAL INFORMATION:
; APPLICANT: MISHRA, LOPA
; TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT. . .
; FILE REFERENCE: XX/P04470US0
; CURRENT APPLICATION NUMBER: US/08/841,349B
; CURRENT FILING DATE: 1997-04-30
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2154
; TYPE: PRT
; ORGANISM: Mus musculus
; US-08-841-349-4

Query Match          5.3%; Score 291.5; DB 1; Length 2154;
Best Local Similarity 20.3%; Pred. No. 2,8e-10;
Matches 194; Conservative 146; Mismatches 326; Indels 291; Gaps 38;

Qy      287 KPNASQSLVWCKEVTKNYKGVKITNFTTSWRNGLSFCAILHFRPDLIDYKSLNPDQIK 346
Db      159 KKSAXDALLMCMQMTAGYPNVVIHFTTSKRDGMAFNLLIHGRPDLIDPDCLKSSNAH 218
Qy      347 ENNKAYD-GFASIGISRLLEPSDMLLAI PDKLTVTYLYLQIDRAHFGQELNVQ---- 401
Db      219 YNLQNAFNLAEOHGLTKLDPED-LSVDHPDEKSIITVYVYHYFSKKKALAVEGKI 277
Qy      402 --IEENSSKSTYKVGNYET-----DTNSSV 424
Db      278 GKVLDNAITEKKEIKETLADSLDEWIBQTIILNNRFRANSLVGVOOQLAFTYRTV 337
Qy      425 DQKFAEISDLK-----RBPLOQPIGAVDFLSQDDSVFVND 464

```



```

Db      338  EKPPEKTEKGNLEVLPAIOSKRRANQKVMPEBKILSDINKAMERLEK----- 388
Qy      465  GVGESESEHQTPDDHLSPTASTPYCRKTSPTBEPKSOOSSR-----TSSDDPGICS 518
Db      389  --AEHERELALNNELIROEKLSQLARFRKAMERETWLSNQRLVSQDNFGFDLPAVEA 446
Qy      519  NTDSQA-----QVLLGKKRLKAKT-----LEL----- 542
Db      447  ATKKEALETDIAAVERVQAVVAARELEANNYDIKRITARKDNVIRLMEYLELRA 506
Qy      543  -----SDLYSDKKKDMSPPICB-----TDEOKLQTLIDIGS 575
Db      507  RRORLEMLGLQKIFQEMLYIMDMWDEMKVLLSQDYGHLLGVEDLLQKALVEADLAI 566
Qy      576  NLEKKELENSRLSECRSDE--SPIKTSL--SPTSKLGYSRDLDAKKKHAIRQT-- 630
Db      567  QAEVRGVNAAQKAPATGEGYKPCDPVIRDRVAHMEFCYQELCOLAERARLEBSRR 626
Qy      631  -----ESDPDADRTTLNHAHDSKIVQ--HRLLSR-----OEBLEKARVLLBOAR 674
Db      627  LMKPFWEABEEMGIRREKEXILSSDDYGKDLTSVRLSKRAFEDEMSGSG--HFBOAI 685
Qy      675  RDAALKAGNKNTATPFCNRLSDQDEERRRLRERARQLIAEASGVWMSLPYSG 734
Db      686  KEGEDMIAEEN-----FGSEKIRERIIYIRQMANL--EQLSAIRKKRLER-- 729
Qy      735  EMAAEKLKERSKASGDENDNIEIDTNEBIPGFFVGGD-----ELTNL 778
Db      730  ---ASLHQFOADADDIDAMMLDILKIVSSNDV--GHDEVSTQSLVKKHDAVEIRTN 783
Qy      779  ENDLDT--PEONSKIVDLKAKLLEVOPOV--ANSFSSAAQKAVTSSBODMSGTEDLATE 836
Db      784  RPTIDTLHEQASAL-----POAHASDPVKGRLAIGERCCEMAELT--LKQ 830
Qy      837  RLQKTERFRNPFVFSKOSTVAKTQLOSPQYIENRPEMKQRSIOEDTKKNEKAIT 896
Db      831  ALQDTLALYK--MFS-----EADACELMIDKEQMLNMQIP----- 866
Qy      897  ETORRPSDEVLANKFKQTSQYVGBLALENEQOQIDTPRALVEKRLRYLMDTGRNTE 956
Db      867  ---KLDELLEVIQHRFB-----SLEPENNOASRAVAVNOIARQLMHNHGPSEK 911
Qy      957  EEAAMQO-----WF-----MLVNKKAALIRMNQSLSEKHNLEBRYELNRELAMLA 1007
Db      912  EIRAOQDKLNTRWQOFRELVDRKKDAL--LSALSI-----ONYHLECNETKSCIR- 959
Qy      1008  EDWQTEAQRREQLLD--ELVALVNRDALVRLDAOE-----KOAE--EDEH 1054
Db      960  ---EKTIVIESQDLGNDLAGVMALOCKLTGMERDLVAIEAKLSLDQKAEKLESEH 1013

```

## RESULT 7

US-09-431-184A-4.

; Sequence 4, Application us/09431184A

; Patent No. 6642362

; GENERAL INFORMATION:

; APPLICANT: MISHA, LOEA

; TITLE OF INVENTION: GENES CODING FOR EARLY LIVER DEVELOPMENT...

; FILE REFERENCE: P04470US1/BAS

; CURRENT APPLICATION NUMBER: US/09/431,184A

; CURRENT FILING DATE: 1999-11-01

; PRIOR APPLICATION NUMBER: PCT/US98/08656

; PRIOR FILING DATE: 1998-04-30

; PRIOR APPLICATION NUMBER: US 08/841,349

; PRIOR FILING DATE: 1997-04-30

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 4

; LENGTH: 2154

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-431-184A-4

Query Match 5.3%; Score 291.5; DB 2; Length 2154;  
 Best Local Similarity 20.3%; Pred. No. 2.8e-10;  
 Matches: 194; Conservative 146; Mismatches 326; Indels 291; Gaps 38;

```

Qy      287  KENASQSLVWCKEVTKNYRQVKTITFTTSNRNGLSFCAIIHFRPDLIDYKSLNPQDIK 346
Db      159  KKSADALLMWCQMKTAGYPRVNIHNFFTSMDGAFAALIHKRPDLIDPDKLKSNAH 218
Qy      347  ENNKAYD--GFAISGISLLEPSPMTLALIDKLTWMTYLYQIRAHFSGQELANVQ---- 401
Db      219  YNLQNAFMAIOHGLTGLDLPED--ISVDHDEKSIITYVTVYNYFESKMAANEGRI 277
Qy      402  --IENSSKSYKYGNVT-----DTNSSV 424
Db      278  GKVDNALIETKMTIKETTLASDLLEWIBQTIILNKRKANSLVGVOOQLQAFNTRYV 337
Qy      425  DOEKFYAEISDLK-----REPELOPISGAVDFLSODDSVFVND 464
Db      338  EKPPEKTEKGNLEVLPAIOSKRRANQKVMPEBKILSDINKAMERLEK----- 388
Qy      465  GVGESESEHQTPDDHLSPTASTPYCRKTSPTBEPKSOOSSR-----TSSDDPGICS 518
Db      389  --AEHERELALNNELIROEKLSQLARFRKAMERETWLSNQRLVSQDNFGFDLPAVEA 446
Qy      519  NTDSQA-----QVLLGKKRLKAKT-----LEL----- 542
Db      447  ATKKEALETDIAAVERVQAVVAARELEANNYDIKRITARKDNVIRLMEYLELRA 506
Qy      543  -----SDLYSDKKKDMSPPICB-----TDEOKLQTLIDIGS 575
Db      507  RRORLEMLGLQKIFQEMLYIMDMWDEMKVLLSQDYGHLLGVEDLLQKALVEADLAI 566
Qy      576  NLEKKELENSRLSECRSDE--SPIKTSL--SPTSKLGYSRDLDAKKKHAIRQT-- 630
Db      567  QAEVRGVNAAQKAPATGEGYKPCDPVIRDRVAHMEFCYQELCOLAERARLEBSRR 626
Qy      631  -----ESDPDADRTTLNHAHDSKIVQ--HRLLSR-----OEBLEKARVLLBOAR 674
Db      627  LMKPFWEABEEMGIRREKEXILSSDDYGKDLTSVRLSKRAFEDEMSGSG--HFBOAI 685
Qy      675  RDAALKAGNKNTATPFCNRLSDQDEERRRLRERARQLIAEASGVWMSLPYSG 734
Db      686  KEGEDMIAEEN-----FGSEKIRERIIYIRQMANL--EQLSAIRKKRLER-- 729
Qy      735  EMAAEKLKERSKASGDENDNIEIDTNEBIPGFFVGGD-----ELTNL 778
Db      730  ---ASLHQFOADADDIDAMMLDILKIVSSNDV--GHDEVSTQSLVKKHDAVEIRTN 783
Qy      779  ENDLDT--PEONSKIVDLKAKLLEVOPOV--ANSFSSAAQKAVTSSBODMSGTEDLATE 836
Db      784  RPTIDTLHEQASAL-----POAHASDPVKGRLAIGERCCEMAELT--LKQ 830
Qy      837  RLQKTERFRNPFVFSKOSTVAKTQLOSPQYIENRPEMKQRSIOEDTKKNEKAIT 896
Db      831  ALQDTLALYK--MFS-----EADACELMIDKEQMLNMQIP----- 866
Qy      897  ETORRPSDEVLANKFKQTSQYVGBLALENEQOQIDTPRALVEKRLRYLMDTGRNTE 956
Db      867  ---KLDELLEVIQHRFB-----SLEPENNOASRAVAVNOIARQLMHNHGPSEK 911
Qy      957  EEAAMQO-----WF-----MLVNKKAALIRMNQSLSEKHNLEBRYELNRELAMLA 1007
Db      912  EIRAOQDKLNTRWQOFRELVDRKKDAL--LSALSI-----ONYHLECNETKSCIR- 959
Qy      1008  EDWQTEAQRREQLLD--ELVALVNRDALVRLDAOE-----KOAE--EDEH 1054
Db      960  ---EKTIVIESQDLGNDLAGVMALOCKLTGMERDLVAIEAKLSLDQKAEKLESEH 1013

```

## RESULT 8

US-09-538-092-1243

; Sequence 1243, Application us/09538092



333 LANS-----LSGVNQLOSPNSYRT--VEKPP--KTEKGNL 365  
QY 526 QVL-----GKKRLKAELELSDLYSDKKOMSPPIGCEDEQLOTLDIGS 575  
Db 366 EVLFTTISKLRANNQKTYTPREGRLISDINKAMRLERA-----EHERELAKT---E 416  
QY 576 NLEKELNSRLSECRDPESEPIKTSPTSCLKGYSTRDLAK-----KGNASLQOT 630  
Db 417 LIRQKLE---QLARFPRKAMRETWLSNQRLVSQNFQLELAABAARKHEAL-ET 472  
QY 631 ESDPADTTLNADHSHKLYOH---RLSRQEE-----LKE---RATYLL--- 670  
Db 473 DIVAVSGVQAVDAVAELAHEHYDIKRIARQNNVAVLWDFLEEMVAARERILLME 532  
QY 671 -----EQARDAALAKGNKHN- 688  
Db 533 LQKVPQDLLYMDMAEKGRLOQODLKLHAGVEDLLQLHLEVADIAVOAERVAVSA 592  
QY 689 TATPCN-----ROLSDQDEER---RQLEBARQLIAEARSQVMSG---LPSYGE 735  
Db 593 SALRFCDGKGYRPGCPQVSEBVALTLEQSYEALCELAATRRARLEBRRLMFLMEYGE 652  
QY 736 MAEKLEKRSASGDENDNIBITNEKIPBGFVVGSDLTNLENDL---TPEQNSK 790  
Db 653 AEAWVREQOHLASLET-----GRDLTGVLRLNKHMTALRBMGSR 693  
QY 791 LVDLKLKLEVOQVANSPPSAQKAVTESSE--ODMSTGEDLATERLOCTTERFRPV 849  
Db 694 LGPKL---TLEQGOQLVAEGHPGANOASTRAALQAOERLEBALEERQAOLAQ--ASLY 750  
QY 850 VFSKDSIVTKTQLOFSQYIENRPEM-----KRSIOEDTKGNEKALITET 898  
Db 751 QFOADANMEMAVDALRLVSS--PEVGHDEPSTQALAQHRALEBETIAAHPTLDALAE- 808  
QY 899 QKQSEDEVLNKGFQDTSQYVGBELALENEQKQIDTRALVEKRLRYLMDTGRTSEEE 958  
Db 809 -----QAALPRLASHTPE--VQGRVPTLEQHYEELQABA-----GERARALE 849  
QY 959 AMQEWPL-----VKNKNALIRMNQLOSLLEKHD---LEERYELNRLERATLA 1006  
Db 850 AALPYTTLSEAGACGLWBEKE---QWNLGLALPERLEDEBVVQORETLEPEMNALAA 906  
QY 1007 IEDMQTEAQKRRRLT-----LDELVAL---VKNRDALVRDLDAQKQK---BEED 1052  
Db 907 ----RITAVSDIAQQLKASPPGKDRITIGTQEQALNQRFQFRSLAGGKALTSALSIQN 962  
QY 1053 EHLERTLEQ 1061  
Db 963 YHLECTETQ 971

RESULT 10  
US-09-949-016-6777  
Sequence 6777, Application US/09949016  
Patent No. 681239  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYOMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CLO01307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FASTSEQ for Windows Version 4.0  
SEQ ID NO 6777  
LENGTH: 1026  
TYPE: PRT

ORGANISM: Human  
US-09-949-016-6777  
Query Match 5.2%; Score 285; DB 2; Length 1026;  
Best Local Similarity 21.2%; Pred. No. 2.7e-10;  
Matches 242; Conservative 153; Mismatches 363; Indels 384; Gaps 58;  
QY 116 DDP---DAEELNPFQDP-----DSEBRTETASPRTEDESFNYSNYPPEVOTPOLY 165  
Db 79 DGRPSDSVEGSPFRPSPSFAVDEDEKFLASGCT-----YLDNDNIELVDTPQTL 130  
QY 166 NP-----FD-----EPEA 173  
Db 131 EPRASDAKNGQKVNTRKSTDSVPISKSTLSRLSLQASFPDQASSGNPEAVALARDA 190  
QY 174 FVTIKDSPPOSTKR-KNIRPYDMSKYLYADSSKTEBEBELDSNPPEKSTPPNNLY- 230  
Db 191 YSTGSSSASSTLKTKKPRPPLKCK-----KOTTKKPTETPPVKETQGFDEBSLVPS 243  
QY 231 --NPVQLETERRYKRAKAP-----PVLSPKT-----QVLN-----E 261  
Db 244 GENLASSETKIE-SAKTBGSPFALLIETPLBPAVGKACPLDSASABGVPPASGGGRVQ 302  
QY 262 NTVSAK---DLSTSPKSPPI-PSPVLGRKENASQSLVWCKEYTKNYRGYKITYFTSM 317  
Db 303 NSPPVGKRTPLPLTAPBAGEVTPSDSGQEDSPAKGLSV---RLBFQDSBK-----SSW 354  
QY 318 RNLGSPCALHFRPDLIDYKSLNPQDIKENNKAYVGFASIGISRLIEPSDMVLAIPD 377  
Db 355 -----DNOQENPPPTKIGKK-----PVAKMPLRPRK---MKKTPE 387  
QY 378 KLTWTYLYQIRAHPSQGLNVQIENSSKSTYKVGNYETDNTSSVQGEK--YAEISD 435  
Db 388 KLD-NTPASPR---SPABNDIPY---AKGTY---TFDDI---KWDENPNPSSSTSK 433  
QY 436 LKREPELQPISGAVFLSDPDSVFVNDSGVSESESHQTPDDHLSPESTASPYCGRKSD 495  
Db 434 MQBSPKLPQ-----QSNFPDD--TCBSVDPFTSSKTP 466  
QY 496 TEPOKS-----QOSSGRTSGSDPGICSNDSIOAOVLAKKRLKAELELSDLYSD 549  
Db 467 SSPKSPASFEIPIPSAMEANGVDGGL--NKPARK-----KKTPLKTDTRVVK---SP 515  
QY 550 KKKMSP-----FICERTDOKL-----QTLDSGSLBEKEL 582  
Db 516 KRSPLSDPPSQDPPTAETPETPVISAVVHAHTDEKLAVTNQKTCMTVD---LEADKQ 571  
QY 583 ENSRLSECRSDPESPIKTSI-SPTSKLGYSYSRDLDAKKGASLRQTESDPADRTTL 641  
Db 572 DYPOP---SDLSTFVNETKSSFTBELDYRNSYIETMEKIGSSLFODDAP----- 620  
QY 642 NHADHSHKLYOHRLLSQEBELKERARVYLBOARRDALKAGNKNTATPFCNRQLSDQ 701  
Db 621 -----KQALYIMFTQSGSPVKSFPVMSSPPT-----CSGS 654  
QY 702 ODEERRQLERA-----ROLIAEARSQVMSSELVSYGMALAKKESKASGDERDN 754  
Db 655 SFETELAVTVAANOHPVPRGLAPNOESHLOVEKESQKELKEMGL-----GTSEEA 707  
QY 755 IEIDTNEIPEGFVVGDELITNLNDL-----TPPEONSKLVDLKLLKGLLEVQ 803  
Db 708 IEITA---PEGSYASDALISRLAHVSLGALDYLEPDLAEKONPPLFAOKLOBELEFA 763  
QY 804 PAVANSPPSAQAQVATSSSQDMKSGTEDLATERLOKT-----TERFRNP--VVSFK- 853  
Db 764 IMRIEALKARQIALASRSHQDAKRAAHPTDVSISKTALYSRIGTAVEKFPAGILFOQP 823  
QY 854 --DSTVAKTQLOGSS-----QYENRPEKRRQSGOEDTKKNEEALATTEQOK 901  
Db 824 DLDSALQIAREIITTKEREVSEWMDKYESSRREVENRKIVAEYK-----TTAQ 873  
QY 902 PSEDEVLNKGFQDTSQYVGBELALENEQKQIDTRALVEKRLRYLMDTGRTSEBRAM 961

```

Db      874 MIEDE---QREKSVSHQTVQOL-VLEKEQALADLNS--VEKS---LADLFRYERKMEVL 924
Qy      962 QEWFMVLVKNKALIRRMNO--LSLEKEHDLERYELL-----NRELAM--LA 1006
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      925 EGF----RKBEVYLKRCQETLSNVKKE--EQRVQALKVAEKLDRANAELIQVNGKA 977
Qy      1007 IEDMQKTEAQKRRRLDLDELVALVNRDALVRDLDAQEQAEBDEHLERTLEQNGKM 1066
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      978 QOEQAHAQASLRKEQL-----RYDALERTLEQNKKEIEE---LTRICDELIAKM 1023
Qy      1067 AK 1068
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1024 GK 1025

```

```

RESULT 11
US-09-949-016-10870
; Sequence 10870, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10870
; LENGTH: 1034
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10870

```

```

Query Match      5.2%; Score 285; DB 2; Length 1034;
Best Local Similarity 21.2%; Pred. No. 2,7e-10;
Matches 242; Conservative 153; Mismatches 363; Indels 384; Gaps 58;

Qy      116 DDP--DAEELNPFCDP-----DSEEPITETASPKTEDSPFNNSYMPKCVQTOYL 165
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      87 DGRSRSDVESGPRPSPHSVAIVEDKPIASSGT-----YNLDFDNIELVDTQT 138
Qy      166 NP-----FD-----EPEA 173
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      139 EPRASDAKNOEGKVNTRRKSTDSVPSKSTLSRLSLQASDFDQASSGNPBAVALAPDA 198
Qy      174 FVTIKSPSPQSTGR-KNIRVVDMSKITYADSSKTEEBELDESMPFYPKTPPPNNIV-- 230
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      199 YSTGSSASASTLRTKTKPRPSPSLK-----KQTTKPTETPVKTKQEPDEESLVP 251
Qy      231 --NPVDELTERRYKRAKAP-----PVLSPKT-----GVLN-----E 261
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      252 GEMLASSETKTE-SAKTGGPSPALLERTPLBPAYGPKAACPLDSDSAGVVPSPSGGRVQ 310
Qy      262 NTVSACK--DLSTSPKSPPI-PSPVLAGKPKNMSQSLLVWCKEVTKNYRGVKITNTFTSW 317
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      311 NSPPVGRKTLPLTTAPAGAVTSPDSGGDSPAKGLSV--RLDFPYSBK-----SSW 362
Qy      318 RNLGSLFCALIHRRPDLIDYKSLNPQDIKENNKKAYGFGASIGISRLLESIDMVLAI 377
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      363 -----DMQENPPTPKIKGK-----PVAKMPLRRK--MKKTPS 395
Qy      378 KLTVMYTYLQIRAHFSGQLNVVQIBNSSKSTYKVGNYEDTNSSVDQKPF--YAEISD 435
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      396 KDL-NTPASPR--SPAENDIPF-----AKGTY--TFPID--KWDDENFNPFSSTSK 441
Qy      436 LKKEPELQPIGAVDFLSQDDSVFVNDSGVGESESHQTPDDHLSPSTASPYCRRTKSD 495

```

```

Db      442 MQESPRLPQ-----QSYNFPDP--TCDESVDPFKTSKTP 474
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Qy      496 TEPQKS-----QSSGRTSGDDPGICSNMTDSQAVLLGKRLLAARLEISDLYSD 549
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      475 SSPKSPASFEIIPASAMEANGVDGGL--NKPAK-----KTPPLTDTFPRVK--SP 523
Qy      550 KKKDMSP-----FICETDEOKL-----QTLIGSLLEKEKL 582
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      524 KRSLSDPPSODPTPATPTETPPVISAIVNATBEKLAIVNQKTCMTVD---LEADKQ 579
Qy      583 ENSRSLECRSDPSPPIKTSL-SPTSKLGYSSRDLDLAKKQNASLQTSDDPADRTTL 641
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      580 DYPOP-----SDLSTFVNETKFSSTPELTDYRNSYEIRYEMKIGSSLPODDAP----- 628
Qy      642 MHADHSKIYQNHLLSQEELKERARVLEQARRDALYKANGNTTATPFCRQLSDQ 701
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      629 -----KKQALYLMPTDSQESPVKSPVPMSSPTP-----CSGS 662
Qy      702 QDEERRRQLERER-----ROLIAEARGVKMSLEPSYGEAAAEKLERSKASGDENDN 754
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      663 SPFETELVNTAAKNQHPVVRGLAPNQESHLOVPKSSQKELAMGL-----GTSEA 715
Qy      755 IEIOTNEEIPGFGVVGGDGLTNLENDL-----DTPQNSKLVLDLKLKGLLEVQ 803
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      716 IEIITA-----PEGSFASADALLSRLAHVSLGALDYLEPLAEKNPPLFAQKLQEBLEFA 771
Qy      804 PVAANSFSAQAQAVTRESSQDKKSGTEDLRTERLOKT-----TERFRNP--VVPK- 853
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      772 IMRIEALKLARQIALSRSHQDAKREAAHPDVSISITALLYSRIGTAEVEKPGLLFOQP 831
Qy      854 --DSTVAKTQLOSFS-----QYENRPMKQORSIOEPTKKGNEBKAATTEQOK 901
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      832 DLBSAQIARABITITERBVSWMKDYEBRSRREVMERKLVABYEK-----TIAQ 881
Qy      902 PSEDEVLNKGPKDTSOYVVELAALENEQKQIDTRALYERKLYIMDTGRNTEEBRAM 961
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      882 MIEDE---QREKSVSHQTVQOL-VLEKEQALADLNS--VEKS---LADLFRYERKMEVL 932
Qy      962 QEWFMVLVKNKALIRRMNO--LSLEKEHDLERYELL-----NRELAM--LA 1006
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      933 EGF----RKBEVYLKRCQETLSNVKKE--EQRVQALKVAEKLDRANAELIQVNGKA 985
Qy      1007 IEDMQKTEAQKRRRLDLDELVALVNRDALVRDLDAQEQAEBDEHLERTLEQNGKM 1066
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      986 QOEQAHAQASLRKEQL-----RYDALERTLEQNKKEIEE---LTRICDELIAKM 1031
Qy      1067 AK 1068
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      1032 GK 1033

```

```

RESULT 12
US-09-270-767-42273
; Sequence 42273, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 42273
; LENGTH: 153
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-42273

Query Match      5.2%; Score 283.5; DB 2; Length 153;
Best Local Similarity 53.1%; Pred. No. 2.6e-11;
Matches 52; Conservative 18; Mismatches 27; Indels 1; Gaps 1;

```



```

Db      511 PFNNMMSAMEDLDQMFIVHTIEBI-EG-LISAHDOFSQTLPDADRRERAILAIHKCAOR 568
Qy      799 LLE-----VOQVANSPPSSAOKAVTES-----SQDQKSGEDRTA--- 836
Db      569 IASBNHKLKSGNSPYTTVTQILINSKWEKVQOLVPKDHALLEQSQSQSNHRLRQFAS 628
Qy      837 -----RLQKTERFRNPVFSKSTVRKTQLOSFQY-----IENPEM-----KROSI 881
Db      629 QANVVGWIGTKHEBI-GRISIENKGL-EDQLSHLQYERSIYDYKFNLDLEQCHOLI 686
Qy      882 QE-----DTKKN-----EKKAITETORRPSSEDEVLNKGFKDTQYVVGELALE 927
Db      687 QEALIFNKHNTYMEHIVGMEQLTTIARTINEVENQILTRDAKISQEOQEPFAS 746
Qy      928 NEQKQIDTRALVKEKRLRYLMDGRNTEBE---PAMQEWPMLVNKKNA 973
Db      747 NHFDKQHGALGPEEFACILISGLDYVENDRQGEAEFRIMSLVDPMHS 795

```

## RESULT 15

```

US-09-091-501B-8
; Sequence 8, Application US/09091501B
; Patent No. 6518413
; GENERAL INFORMATION:
; APPLICANT: Tinsley, Jonathan M
; APPLICANT: Davies, Kay E
; TITLE OR INVENTION: utrophin gene expression
; FILE REFERENCE: 620-42
; CURRENT APPLICATION NUMBER: US/09/091.501B
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: PCT/GB96/03156
; PRIOR FILING DATE: 1996-12-19
; PRIOR APPLICATION NUMBER: GB 9525962.8
; PRIOR FILING DATE: 1995-12-19
; PRIOR APPLICATION NUMBER: GB 9615797.9
; PRIOR FILING DATE: 1996-07-26
; PRIOR APPLICATION NUMBER: GB 9622174.2
; PRIOR FILING DATE: 1996-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 8
; LENGTH: 2008
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (239) ... (250)
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric; Xaa = unknown
US-09-091-501B-8

```

```

Query Match      4.98; Score 270.5; DB 2; Length 2008;
Best Local Similarity 20.68; Pred. No. 6.1e-09;
Matches 211; Conservative 164; Mismatches 349; Indels 301; Gaps 53;

```

```

Qy      286 RKNPASQSLVWCKEVTKNRGVKITFTTSMRNGSLFCAILHHFRPDLIDYSLNPDQI 345
Db      148 QQTNSKILISWVRQTRPFSQVAVNLFTTSWIDGLAFNAVLHKKKDLFSMDRVVMS 207
Qy      346 KENNKAAYD-GFASIGISRLLEPSDNV-----LAIAPDKLT-----VM 382
Db      208 IERLEHAFSAHYTLGIEKLLDPEDVAVHL.PXXXXXXXVYEVLPQGVTTDAIREVE 267
Qy      383 TYIYQIRAHSGOELNV---VQIEBNSKSTYKGVNTEDTNNSVDOEKFYAELSLKRE 439
Db      268 TLPRKTKKECEBEIINQSAVLAEBSQP-----RAETPSTVTEVMDMLDSYQIAL 319
Qy      440 PELQOPISGAVDPLSODP-----SVFVNDSGVGE----- 468
Db      320 EVLTWILSLADTQEOQDDIDDDVEVKEQPATHETFMELTAHQSSVGSVLAQANQMLQ 379
Qy      469 ---SESEHQTPDDHLSFSTASPYCRRTKS-----DTEPOKSSQSSG----- 506

```

```

Db      380 GTISEEEFEIQRQMTLLNARWEALRVESMEROSRLDALMLEQKQLOQLSSWLTALTE 439
Qy      507 RTSGSDDPGICSNWDSQAOVNLGKKRL---LKAETLESLD-----Y 546
Db      440 RIQKESPPPLGDDLPSTIQ-KLQBNHKSLONDLEAOVKVNSLTHMVIVDENSESATAL 498
Qy      547 VSDKQKMSPPF--ICEETDEQ--KLQTLDIGSNLEKEKLENSRSLSEC-RSDPESPIKKT 601
Db      499 LEDQQLGLGRMTAVACWMTERRNRIQEISI---LWQELLEBOCLLEAWLTBEKBALNV 555
Qy      602 SLSPSTKLGYSYRDLDLAKKQIASLRQTESPPDADRTTLLNHADHSKIYQ----- 652
Db      556 QTS-----NFKDQKELSGSVRLALIKE---DHMKRQTL---DQISEIQDVQGLSNP 604
Qy      653 ---HLLSROELKEBRRVLEQAR-----RDAALKAG----- 682
Db      605 KASKKMSDSEELTORWDSLVQRLBESSNOVTOVAALGASQIPQDQDLETVHREKGV 664
Qy      683 NKNNTWTATPFCNRQLSDQODE-----ERRQLREARQL--IAEASGVMSLPSYGE 735
Db      665 KKKQKQELPPLTGAHMAQKRSITTELGENTQELRLDQKEMVNAKLMINRTLE-----E 720
Qy      736 MAAEK---LKERKASGDENDNIEIDTNE--BIPEGVVGGDDELTNLENDLDTPEQNS 789
Db      721 MLSCKSLSLPERDKIS-ESLRTVMTMWNKICREVP-----TTLKECIQEPSSVS 768
Qy      790 -----KLVDLKLKQLLEVQOVANSPPSSA--OKAVTES-----RQDKSGT 830
Db      769 QTRIAHPNVQKVLVSSASDIPVQSHRTSEISIPADLDITTELMDVLIDQMLKSN 828
Qy      831 EDL-RTERLQKTERFRNPVFSKSTVRKTQLO---SFSQYIENR---PEMK----- 876
Db      829 VYGVDBEINKIYSRK---ITRADLEQNHQPDYFTLAQNLKXKASSDMKTAITEKL 885
Qy      877 -----RQNSIQE---DTKKNKEKAITETORK-----PSEDEV 908
Db      886 ERYVGNQMDGTQHGVEIRQQLLEMIIDSLQMDHREBTEELMKYEALTYILQARDDL 945
Qy      909 NKGFQTSQYVVGELALBENQOIPTRALVKEKL-RYLMDSGRNTEE-EZAMQEWPM 966
Db      946 TKQISD-NQILLQELGSGEIVNAPDN---VLQKLEBYSQSDTRNVKETEYLTQTSW-- 999
Qy      967 LVNKKNALIRRMNQLSLLEKEHDLERRVYELNRELPAMLAIEDM-OKTEA-----QKR 1018
Db      1000 -IMKQSIADRON--ALBAEW---RTVQASRRDLNPL--KMIQEAEFTVAVLVASH 1049
Qy      1019 REQLLDELVALVYKEDALVRDIDAOEKQAEER-DEH--LERTLEQNGKMAK---KEE 1071
Db      1050 RENVLQDSILA-----RELKQOMQODIQAEIIDAHNDFKSIDGNRQRMVYALGNSSE 1100
Qy      1072 KCVLQ 1076
Db      1101 ATMLQ 1105

```

```

Search completed: May 20, 2006, 17:16:29
Job time : 60 secs

```

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using bw model

Run on: May 20, 2006, 17:26:58 ; Search time 186 Seconds  
(without alignments)  
2679.673 Million cell updates/sec

Title: US-09-856-723A-8

Perfect score: 5489

Sequence: 1 MKQYASMPQTQDVKLFKFP.....RTLEHNGKMAKKEKCVLQ 1076

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2097797 seqs, 463214858 residues

Total number of hits satisfying chosen parameters: 2097797

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*  
2: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*  
3: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*  
4: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US10A\_PUBCOMB.pep:\*  
5: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US10B\_PUBCOMB.pep:\*  
6: /EMC\_Celerra\_SIDS3/ptodata/2/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5486	99.9	1076	4 US-10-275-595A-26	Sequence 26, Appl
2	5173	94.2	1023	5 US-10-450-763-53242	Sequence 5342, A
3	4485	81.7	882	4 US-10-408-765A-1942	Sequence 1942, Ap
4	3151	57.4	669	5 US-10-450-763-53241	Sequence 53241, A
5	1045	19.0	221	5 US-10-450-763-53239	Sequence 53239, A
6	1020.5	18.6	957	6 US-11-097-143-39276	Sequence 39276, A
7	989	18.0	479	5 US-10-450-763-53240	Sequence 53240, A
8	945.5	17.2	1523	4 US-10-473-576-5	Sequence 5
9	897	16.3	175	3 US-09-864-761-42978	Sequence 42978, A
10	897	16.3	175	4 US-10-028-386-33821	Sequence 33821, A
11	510.5	9.3	996	6 US-11-097-143-23565	Sequence 23565, A
12	507.5	9.2	1010	4 US-10-359-012-18	Sequence 18, Appl
13	498.5	9.1	863	4 US-10-359-012-2	Sequence 2, Appl
14	498.5	9.1	863	4 US-10-359-012-23	Sequence 23, Appl
15	498.5	9.1	863	6 US-11-169-041-167	Sequence 14, Appl
16	461	8.4	91	3 US-09-864-761-34415	Sequence 34415, A
17	461	8.4	91	3 US-09-864-761-46554	Sequence 46554, A
18	441	8.0	1480	4 US-10-359-012-22	Sequence 22, Appl
19	431	7.9	1026	4 US-10-359-012-23	Sequence 23, Appl
20	429.5	7.8	904	4 US-10-359-012-16	Sequence 16, Appl
21	418	7.6	79	3 US-09-864-761-37001	Sequence 37001, A
22	418	7.6	79	3 US-09-864-761-46885	Sequence 46885, A
23	399.5	7.3	1633	4 US-10-359-012-4	Sequence 4, Appl
24	368.5	6.7	1004	4 US-10-291-172-235	Sequence 235, App
25	368.5	6.7	1004	4 US-10-221-278-235	Sequence 6, Appl
26	368.5	6.7	1965	4 US-10-359-012-6	Sequence 611, App
27	366.5	6.7	1025	4 US-10-291-172-611	Sequence 611, App

28	366.5	6.7	1025	4 US-10-221-278-611	Sequence 611, App
29	333	6.1	791	4 US-10-170-385-57	Sequence 57, Appl
30	327.5	6.0	1452	5 US-10-995-561-778	Sequence 778, Appl
31	326.5	5.9	3433	4 US-10-408-765A-731	Sequence 731, Appl
32	326.5	5.9	3433	4 US-10-408-765A-732	Sequence 732, Appl
33	326.5	5.9	3433	5 US-10-756-149-5681	Sequence 5681, Ap
34	319	5.8	65	4 US-10-425-115-333910	Sequence 333910,
35	319	5.8	1790	5 US-10-732-923-3332	Sequence 3331, Ap
36	318.5	5.8	1268	5 US-10-732-923-3332	Sequence 3332, Ap
37	318.5	5.8	1790	4 US-10-369-493-1586	Sequence 1586, Ap
38	318.5	5.8	1790	5 US-10-732-923-3330	Sequence 3330, Ap
39	316.5	5.8	2291	6 US-11-097-143-12420	Sequence 12420, A
40	316.5	5.8	4101	4 US-10-369-493-6509	Sequence 6509, Ap
41	316.5	5.8	4101	4 US-10-369-493-6510	Sequence 6510, Ap
42	309	5.6	2326	4 US-10-369-493-6374	Sequence 6374, Ap
43	308.5	5.6	2137	5 US-10-756-149-4806	Sequence 4806, Ap
44	307.5	5.6	2106	4 US-10-408-765A-2093	Sequence 2093, Ap
45	305.5	5.6	915	6 US-11-097-143-8538	Sequence 8538, Ap

#### ALIGNMENTS

```
RESULT 1
US-10-275-595A-26
: Sequence 26, Application US/10275595A
: Publication No. US20040078804A1
: GENERAL INFORMATION:
: APPLICANT: YUE, Henry
: APPLICANT: TANG, Y. Tom
: APPLICANT: AU-YOUNG, Janice
: APPLICANT: LU, Dyrung Alma M.
: APPLICANT: BAUGHN, Mariah R.
: APPLICANT: HITLMAN, Jennifer L.
: APPLICANT: AZIMZAI, Valda
: APPLICANT: LAI, Preeti
: APPLICANT: YAO, Montague G.
: APPLICANT: BANDMAN, Olga
: APPLICANT: BURFORD, Neil
: APPLICANT: BATRA, Sayeev
: APPLICANT: KEARNEY, Liam
: APPLICANT: POLICKY, Jennifer L.
: TITLE OF INVENTION: CYTOSKELETON-ASSOCIATED PROTEINS
: FILE REFERENCE: PF-0772 USN
: CURRENT APPLICATION NUMBER: US/10/275,595A
: CURRENT FILING DATE: 2003-06-13
: PRIOR APPLICATION NUMBER: US 60/201,960
: PRIOR FILING DATE: 2000-05-05
: PRIOR APPLICATION NUMBER: US 60/202,729
: PRIOR FILING DATE: 2000-05-08
: PRIOR APPLICATION NUMBER: US 60/209,705
: PRIOR FILING DATE: 2000-06-05
: PRIOR APPLICATION NUMBER: US 60/210,149
: PRIOR FILING DATE: 2000-06-07
: PRIOR APPLICATION NUMBER: US 60/213,215
: PRIOR FILING DATE: 2000-06-21
: NUMBER OF SEQ ID NOS: 68
: SOFTWARE: PERL Program
: SEQ ID NO 26
: LENGTH: 1076
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURES:
: NAME/KEY: misc_feature
: OTHER INFORMATION: Incyte ID No: 5202390CD1
US-10-275-595A-26
Query Match 99.9%; Score 5486; DB 4; Length 1076;
Best Local Similarity 99.9%; Pred. No. 1.9e-287;
Matches 1075; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
1 MKQYASMPQTQDVKLFKFPSSKVVSAALQFSICIFLEGGKATDEDMOSIASVSKQ 60
|||||
```

Db 1 MKQYASPMPTQTDVAKLKFKPLSKVVSAAALQPSLSICIFLREGKATDEDMOSLASIMSMKQ 60  
QY 61 ADIGNLDPEEDNEDDENRNVNOBEKAKITELINKLNFLEBAEKDLATVNSNPPDDDA 120  
Db 61 ADIGNLDPEEDNEDDENRNVNOBEKAKITELINKLNFLEBAEKDLATVNSNPPDDDA 120  
QY 121 AELNPFQDPSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBEAFVTKDS 180  
Db 121 AELNPFQDPSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBEAFVTKDS 180  
QY 181 POSTKRNIRPVMSKYLVAADSKTEEBELDESNPFYEPKSTPPNNLVNPFQELTER 240  
Db 181 POSTKRNIRPVMSKYLVAADSKTEEBELDESNPFYEPKSTPPNNLVNPFQELTER 240  
QY 241 RVKRAKAPAPVLSKPTGVLENENTVSAGKDLSTSPKPSPIPSPVLGKRNASQSLVWCKE 300  
Db 241 RVKRAKAPAPVLSKPTGVLENENTVSAGKDLSTSPKPSPIPSPVLGKRNASQSLVWCKE 300  
QY 301 VTKXNRGVKITNFTTSWRNGLSFCALLHFRPDLIDYKSLNPODIKENNKKAYDGFASIG 360  
Db 301 VTKXNRGVKITNFTTSWRNGLSFCALLHFRPDLIDYKSLNPODIKENNKKAYDGFASIG 360  
QY 361 ISRLBESDMVLAIIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDT 420  
Db 361 ISRLBESDMVLAIIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDT 420  
QY 421 NSGVDEKFAELSDLKREPELOPISGAVDFLSQDSDSVFVNDGVSSESEHQTDPDHL 480  
Db 421 NSGVDEKFAELSDLKREPELOPISGAVDFLSQDSDSVFVNDGVSSESEHQTDPDHL 480  
QY 481 SPSTASPYCARTKSDTPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETL 540  
Db 481 SPSTASPYCARTKSDTPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETL 540  
QY 541 ELSDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBESRSPESPIKK 600  
Db 541 ELSDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBESRSPESPIKK 600  
QY 601 TSLSPSTKLGYSYSRDLDLAKKGAASLROTESPDPADRTLNADHSSKTVORHLSROE 660  
Db 601 TSLSPSTKLGYSYSRDLDLAKKGAASLROTESPDPADRTLNADHSSKTVORHLSROE 660  
QY 661 ELKERARVLLLEQARRDALAKGNKHNNTATPFCNRLSDQODEERRQLREBARQLIAE 720  
Db 661 ELKERARVLLLEQARRDALAKGNKHNNTATPFCNRLSDQODEERRQLREBARQLIAE 720  
QY 721 ARSGVMKSELPSYGEAAAEKLERKSKASGDBNDNIEITNBEIPEGFTVGGDBLTNLEN 780  
Db 721 ARSGVMKSELPSYGEAAAEKLERKSKASGDBNDNIEITNBEIPEGFTVGGDBLTNLEN 780  
QY 781 DLDTPEONSKLVYDKLKKLEVOPOVANSPSAAOKAVTSSSEODMKSGETEDRTERLOK 840  
Db 781 DLDTPEONSKLVYDKLKKLEVOPOVANSPSAAOKAVTSSSEODMKSGETEDRTERLOK 840  
QY 841 TTERFRNPVVFSSKDSYVRKTLQOSFSQYIENRPEMKQORSIOBDYKGNNEKAAITETOR 900  
Db 841 TTERFRNPVVFSSKDSYVRKTLQOSFSQYIENRPEMKQORSIOBDYKGNNEKAAITETOR 900  
QY 901 KPSBEDVLNKGFDOTQYVVGELAALENEKQIDTRALVEKRLYLMTGRNTBEEBAM 960  
Db 901 KPSBEDVLNKGFDOTQYVVGELAALENEKQIDTRALVEKRLYLMTGRNTBEEBAM 960  
QY 961 MOSEMFVLNKNALIRMNQSLLEBKHDERRYELLNRELRLAMLAIEDQKTEAQRRE 1020  
Db 961 MOSEMFVLNKNALIRMNQSLLEBKHDERRYELLNRELRLAMLAIEDQKTEAQRRE 1020  
QY 1021 QLLLDLVALVNRDALVRLDIAOEKQAEEDHLERTLEONKGNKAKKEKCVLQ 1076  
Db 1021 QLLLDLVALVNRDALVRLDIAOEKQAEEDHLERTLEONKGNKAKKEKCVLQ 1076

RESULT 2  
US-10-450-763-53242

/ Sequence 53242, Application US/10450763  
/ Publication No. US20050196754A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Hyseq, Inc  
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
/ FILE REFERENCE: 790CIP3/US  
/ CURRENT APPLICATION NUMBER: US/10/450,763  
/ PRIOR FILING DATE: 2003-06-11  
/ PRIOR APPLICATION NUMBER: PCT/US01/08631  
/ PRIOR FILING DATE: 2001-03-30  
/ PRIOR APPLICATION NUMBER: 09/540,217  
/ PRIOR FILING DATE: 2000-03-31  
/ PRIOR APPLICATION NUMBER: 09/649,167  
/ NUMBER OF SEQ ID NOS: 60736  
/ SOFTWARE: Custom  
/ SEQ ID NO 53242  
/ LENGTH: 1023  
/ TYPE: PRT  
/ ORGANISM: Homo sapiens  
/ FEATURE:  
/ NAME/KEY: DOMAIN  
/ LOCATION: (254)..(284)  
/ OTHER INFORMATION: Actinin-type actin-binding domain proteins domain identified  
/ OTHER INFORMATION: by eMATRIX, accession number BL00019D, p-value=4.200e-19, raw score  
/ OTHER INFORMATION: of 15.33  
/ FEATURE:  
/ NAME/KEY: DOMAIN  
/ LOCATION: (237)..(342)  
/ OTHER INFORMATION: Calponin homology (CH) domain identified by Pfam, accession  
/ OTHER INFORMATION: name CH, E-value=2.4e-17, Pfam score of 71.1  
/ US-10-450-763-53242

Query Match 94.2%; Score 5173; DB 5; Length 1023;  
Best Local Similarity 99.7%; Pred. No. 1,4e-270;  
Matches 1011; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 63 IGNIJDPEEDNEDDENRNVNOBEKAKITELINKLNFLEBAEKDLATVNSNPPDDDAE 122  
Db 10 IGNIJDPEEDNEDDENRNVNOBEKAKITELINKLNFLEBAEKDLATVNSNPPDDDAE 69  
QY 123 LNPFGDPDSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBEAFVTKDSPP 182  
Db 70 LNPFGDPDSEEBITETASPRKTEDSFYNNSYNPFKEVQTPQYLNPFDEBEAFVTKDSPP 129  
QY 183 OSTKRNIRPVMSKYLVAADSKTEEBELDESNPFYEPKSTPPNNLVNPFQELTERRY 242  
Db 130 OSTKRNIRPVMSKYLVAADSKTEEBELDESNPFYEPKSTPPNNLVNPFQELTERRY 189  
QY 243 KRKAPAPVLSKPTGVLENENTVSAGKDLSTSPKPSPIPSPVLGKRNASQSLVWCKEYT 302  
Db 190 KRKAPAPVLSKPTGVLENENTVSAGKDLSTSPKPSPIPSPVLGKRNASQSLVWCKEYT 249  
QY 303 KNYRGVKITNFTTSWRNGLSFCALLHFRPDLIDYKSLNPODIKENNKKAYDGFASIGIS 362  
Db 250 KNYRGVKITNFTTSWRNGLSFCALLHFRPDLIDYKSLNPODIKENNKKAYDGFASIGIS 309  
QY 363 RLLEPSDMVLAIIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDTNS 422  
Db 310 RLLEPSDMVLAIIPDKLTVMYTLVQIRAHFSGQELNVVQIENSSKSTYKGVNRETDTNS 369  
QY 423 SVDOEKFAELSDLKREPELOPISGAVDFLSQDSDSVFVNDGVSSESEHQTDPDHLSP 482  
Db 370 SVDOEKFAELSDLKREPELOPISGAVDFLSQDSDSVFVNDGVSSESEHQTDPDHLSP 429  
QY 483 STASPYCARTKSDTPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETTEL 542  
Db 430 STASPYCARTKSDTPQKSSQSSGRTSGSDPGICSNSTSTQAVYLLGKKRLKAETTEL 489  
QY 543 SLDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBESRSPESPIKKT 602  
Db 490 SLDLYVSDKKKMSPPFICEETDEQKLTLDIGSNLEKELBNSRSLBESRSPESPIKKT 549



QY 603 LSPYTSKLGYSYRDLAKKHAASLRQTESPDADRTTLNHADSSKIVQHRLLSROBEL 662  
 Db 550 LSPYTSKLGYSYRDLAKKHAASLRQTESPDADRTTLNHADSSKIVQHRLLSROBEL 609  
 QY 663 KERARVLLLEQARBDALKAGNKNNTTATPCNROLSDQODEERRROLREARQOLIAAR 722  
 Db 610 KERARVLLLEQARBDALKAGNKNNTTATPCNROLSDQODEERRROLREARQOLIAAR 669  
 QY 723 SGVKSSELPSYGEAABKLERKSKASGDENDNIEIDTNEIIPGPFVVGSGDELTYLLENDL 782  
 Db 670 SGVKSSELPSYGEAABKLERKSKASGDENDNIEIDTNEIIPGPFVVGSGDELTYLLENDL 729  
 QY 783 DTPSONSKLVDLKIKKLELVOPQVANSPPSAQAQVATESSEBQDMKSGTEDLATERLOKTT 842  
 Db 730 DTPSONSKLVDLKIKKLELVOPQVANSPPSAQAQVATESSEBQDMKSGTEDLATERLOKTT 789  
 QY 843 ERFRNPVVFSDSTYKTKQLOSPSOYIENRPEMKQORSIOEDTKKGNKKAALITETOKRP 902  
 Db 790 ERFRNPVVFSDSTYKTKQLOSPSOYIENRPEMKQORSIOEDTKKGNKKAALITETOKRP 849  
 QY 903 SEDEVLANKGFOTSOYVGEIALAENBOKOIDTRALVEKRLRYLMDTGRNTEEBEAMMO 962  
 Db 850 SEDEVLANKGFOTSOYVGEIALAENBOKOIDTRALVEKRLRYLMDTGRNTEEBEAMMO 909  
 QY 963 EWFMLVNNKNAALIRRMNQLSLEKEHDLERRYLELREMLALAIEDWQKTEAOKRRQOL 1022  
 Db 910 EWFMLVNNKNAALIRRMNQLSLEKEHDLERRYLELREMLALAIEDWQKTEAOKRRQOL 969  
 QY 1023 LLDDELVALVNRDVALVRDLDAQOEKAEDEDEHLERTLEQNKGMKAKKEKCVLQ 1076  
 Db 970 LLDDELVALVNRDVALVRDLDAQOEKAEDEDEHLERTLEQNKGMKAKKEKCVLQ 1023

## RESULT 3

US-10-408-765A-1942  
 ; Sequence 1942, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Wernock, Dale E.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; FILE REFERENCE: 660088.465  
 ; CURRENT APPLICATION NUMBER: US/10/408, 765A  
 ; NUMBER OF SEQ ID NOS: 3073  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1942  
 ; LENGTH: 882  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-408-765A-1942

Query Match 81.7%; Score 4485; DB 4; Length 882;  
 Best Local Similarity 100.0%; Pred. No. 1.5e-23;  
 Matches 882; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 MSKTLIYADSSKTEBEELDESNPFFEPKSTPPPNLVNPFVQLETFBRVYKRAAPPVLP 254  
 Db 1 MSKTLIYADSSKTEBEELDESNPFFEPKSTPPPNLVNPFVQLETFBRVYKRAAPPVLP 60  
 QY 255 KTGVLNENTVAGAGDLSLSPKSPSPSPVGLGRKKNASQSLVWCCKEYVQYRGYKTNPT 314  
 Db 61 KTGVLNENTVAGAGDLSLSPKSPSPSPVGLGRKKNASQSLVWCCKEYVQYRGYKTNPT 120  
 QY 315 TSWNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAYDGFASIGISRLLEPSDWLLA 374  
 Db 121 TSWNGLSFCAILHFRPDLIDYKSLNPQDIKENNKAYDGFASIGISRLLEPSDWLLA 180

QY 375 IPDKLTWVTYLYQIRAHFSGGELNVQIENSSKSTYKGVNVEYEDTNNSSVDQEKFYAELS 434  
 Db 181 IPDKLTWVTYLYQIRAHFSGGELNVQIENSSKSTYKGVNVEYEDTNNSSVDQEKFYAELS 240  
 QY 435 DLKRPBELQPIISGAVDFLSODDSVFVNDGVSSESHQTPDDHLSSTASPYCRRTKS 494  
 Db 241 DLKRPBELQPIISGAVDFLSODDSVFVNDGVSSESHQTPDDHLSSTASPYCRRTKS 300  
 QY 495 DTEPQKSGQSSGRSSGSDPGICNTDSTQAVLLGKRLKAEFLLESLDIYVSKKKDM 554  
 Db 301 DTEPQKSGQSSGRSSGSDPGICNTDSTQAVLLGKRLKAEFLLESLDIYVSKKKDM 360  
 QY 555 SPPFICEETDQKQITDIGNLEKEKLENSRSLSECRSDPSPIKTKTSLSPTSRLGYYS 614  
 Db 361 SPPFICEETDQKQITDIGNLEKEKLENSRSLSECRSDPSPIKTKTSLSPTSRLGYYS 420  
 QY 615 RDLDLAKKHAASLRQTESPDADRTTLNHADSSKIVQHRLLSROBELKERARVLLLEQAR 674  
 Db 421 RDLDLAKKHAASLRQTESPDADRTTLNHADSSKIVQHRLLSROBELKERARVLLLEQAR 480  
 QY 675 RDAALKAGNKNNTTATPCNROLSDQODEERRROLREARQOLIAEARSQVMSSELPSYG 734  
 Db 481 RDAALKAGNKNNTTATPCNROLSDQODEERRROLREARQOLIAEARSQVMSSELPSYG 540  
 QY 735 EMAAEKLERKSKASGDENDNIEIDTNEIIPGPFVVGSGDELTYLLENDLTPSONSKLVDL 794  
 Db 541 EMAAEKLERKSKASGDENDNIEIDTNEIIPGPFVVGSGDELTYLLENDLTPSONSKLVDL 600  
 QY 795 KUKKLELVOPQVANSPPSAQAQVATESSEBQDMKSGTEDLATERLOKTTERRFRNPVFSKD 854  
 Db 601 KUKKLELVOPQVANSPPSAQAQVATESSEBQDMKSGTEDLATERLOKTTERRFRNPVFSKD 660  
 QY 855 STYKTKQLOSPSOYIENRPEMKQORSIOEDTKKGNKKAALITETOKRPSDEVLANKGFKD 914  
 Db 661 STYKTKQLOSPSOYIENRPEMKQORSIOEDTKKGNKKAALITETOKRPSDEVLANKGFKD 720  
 QY 915 TSOYVGEIALAENBOKOIDTRALVEKRLRYLMDTGRNTEEBEAMMOEWFMLVNNKNAAL 974  
 Db 721 TSOYVGEIALAENBOKOIDTRALVEKRLRYLMDTGRNTEEBEAMMOEWFMLVNNKNAAL 780  
 QY 975 IRRMNQLSLEKEHDLERRYLELREMLALAIEDWQKTEAOKRRQOLLDDELVALVNR 1034  
 Db 781 IRRMNQLSLEKEHDLERRYLELREMLALAIEDWQKTEAOKRRQOLLDDELVALVNR 840  
 QY 1035 DALVRDLDAQOEKAEDEDEHLERTLEQNKGMKAKKEKCVLQ 1076  
 Db 841 DALVRDLDAQOEKAEDEDEHLERTLEQNKGMKAKKEKCVLQ 882

## RESULT 4

US-10-450-763-53241  
 ; Sequence 53241, Application US/10450763  
 ; Publication No. US20050196754A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 790CIP3/US  
 ; CURRENT APPLICATION NUMBER: US/10/450, 763  
 ; CURRENT FILING DATE: 2003-06-11  
 ; PRIOR APPLICATION NUMBER: PCT/US01/08631  
 ; PRIOR FILING DATE: 2001-03-30  
 ; PRIOR APPLICATION NUMBER: 09/540,217  
 ; PRIOR FILING DATE: 2000-03-31  
 ; PRIOR APPLICATION NUMBER: 09/649,167  
 ; PRIOR FILING DATE: 2000-08-23  
 ; NUMBER OF SEQ ID NOS: 60736  
 ; SOFTWARE: Custom  
 ; SEQ ID NO 53241  
 ; LENGTH: 669  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:

NAME/KEY: DOMAIN  
LOCATION: (61)..(91)  
OTHER INFORMATION: Actinin-type actin-binding domain proteins domain identified  
OTHER INFORMATION: by EMATRIX, accession number BL00019D, p-value=4.200e-19, raw bcc  
OTHER INFORMATION: of 15.33  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (42)..(147)  
OTHER INFORMATION: Calponin homology (CH) domain identified by Pfam, accession  
OTHER INFORMATION: name CH, E-value=2.4e-17, Pfam score of 71.1  
US-10-450-763-53241

Query Match 57.4%; Score 3151; DB 5; Length 669;  
Best Local Similarity 89.5%; Pred. No. 1e-161;  
Matches 634; Conservative 6; Mismatches 12; Indels 56; Gaps 4;

QY 248 APPVLSFKTGVLENTVSAGKLDSTS--PKSPSPSVLGRKPNASGLLWCKEVTKNY 305  
DB 2 AKPYLYKKYG--NQQGVAALPDVPAWAEAPSPSPVLGRKPNASOSLLWCKEVTKNY 59  
QY 306 RGVKTNFTTSMWNGLSFCALIHFRPDLIDYSLANPDIKENNKRAYDGPASTIGISRL 365  
DB 60 RGVKTNFTTSMWNGLSFCALIHFRPDLIDYSLANPDIKENNKRAYDGPASTIGISRL 119  
QY 366 EPSPDMVLAIIPDKLTWTTLYQIRAHFSGOELNVQIEENSSKSTYKGVNVEITDINSVD 425  
DB 120 EPSPDMVLAIIPDKLTWTTLYQIRAHFSGOELNVQIEENSSKSTYKGVNVEITDINSVD 179  
QY 426 QERFVAELSLKREBELQOPISGAVDFLSODDSVFVNDSGVGESESHQTPDDHLSPTA 485  
DB 180 QERFVAELSLKREBELQOPISGAVDFLSODDSVFVNDSGVGESESHQTPDDHLSPTA 239  
QY 486 SPVCRRTKSTPQKSSQSSGRTSGSDPGICSTNTDSTOAVLLGKRLLKAEITLSD 545  
DB 240 SPVCRRTKSTPQKSSQSSGRTSGSDPGICSTNTDSTOAVLLGKRLLKAEITLSD 299  
QY 546 YVSDKKKDMSPFICEETDEQKQTLIDIGSNLEKEKLENSRLECRSDPSPPIKTSLS 605  
DB 300 YVSDKKKDMSPFICEETDEQKQTLIDIGSNLEKEKLENSRLECRSDPSPPIKTSLS 359  
QY 606 TSKLGYSYSDLDLAKKGAHLRQTESDPADRTTLNHADHSSKIYOHRLLSROEELKER 665  
DB 360 TSKLGYSYSDLDLAKKGAHLRQTESDPADRTTLNHADHSSKIY----- 405  
QY 666 ARVLLEARDALAKAGNKNTATATFCRKQLSDQDEERRRLRARAROLIAEARS 725  
DB 406 -----QDEERRRLRARAROLIAEARS 430  
QY 726 KMSLPSEYGEAAAEKLERKASGDENDNIEIDTNEBIPGFVVGSGDELTLNENDLDP 785  
DB 431 KMSLPSEYGEAAAEKLERKASGDENDNIEIDTNEBIPGFVVGSGDELTLNENDLDP 490  
QY 786 EONSKLVDLKLKLEVOQVANSPPSAQAOKAVTESSEDMKSGTEDLRTERLOKTERF 845  
DB 491 EONSKLVDLKLKLEVOQVANSPPSAQAOKAVTESSEDMKSGTEDLRTERLOKTERF 550  
QY 846 RNVVVSSTKSTVKTQUSFSQYIENPEMKRQSIQEDTRKKGEEKRAALTTEROKRSED 905  
DB 551 RNVVVSSTKSTVKTQUSFSQYIENPEMKRQSIQEDTRKKGEEKRAALTTEROKRSED 610  
QY 906 EVLNKPKTSQYVVGELALENEQKIDTRALVERKLRYLMDTGN 953  
DB 611 E---KPKTSQYVVGELALENEQKIDTRALVERKLRYLMDTGN 655

RESULT 5  
US-10-450-763-53239  
Sequence 53239, Application US/10450763  
Publication No. US20050196754A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 790CIP3/US

CURRENT APPLICATION NUMBER: US/10/450,763  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/006631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 53239  
LENGTH: 221  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (65)..(97)  
OTHER INFORMATION: REV protein (anti-repression transactivator protein) domain  
OTHER INFORMATION: identified by EMATRIX, accession number PF00424A, p-value=9.521e-  
OTHER INFORMATION: 09, raw score of 14.34  
US-10-450-763-53239

Query Match 19.0%; Score 1045; DB 5; Length 221;  
Best Local Similarity 84.4%; Pred. No. 1.e-48;  
Matches 217; Conservative 0; Mismatches 4; Indels 36; Gaps 1;

QY 628 ROTESPPDADRTTLNHADHSSKIYOHRLLSROEELKERAVLLEQARRDAALRAGNKNT 687  
DB 1 RNVSSPDDADRTTLNHADHSSKIYOHRLLSROEELKERAVLLEQARRDAALRAGNKNT 60  
QY 688 NTATPPCNRLSDQDEERRRLRARAROLIAEARSQVKNSELPSYGEAAAEKLERKSKA 747  
DB 61 NTATPPCNRLSDQDEERRRLRARAROLIAEARSQVKNSELPSYGEAAAEKLERKSKA 120  
QY 748 GSENDNIEIDTNEBIPGFVVGSGDELTLNENDLDPENSKLVDLKLKLEVOQVA 807  
DB 121 SG-----EONSKLVDLKLKLEVOQVA 144  
QY 808 NSPSSAQAOKAVTSSBODMKSGTEDLRTERLOKTERFRPVVFSKDSYTRKTQLOSFQ 867  
DB 145 NSPSSAQAOKAVTSSBODMKSGTEDLRTERLOKTERFRPVVFSKDSYTRKTQLOSFQ 204  
QY 868 YIENPEMKRQSIQED 884  
DB 205 YIENPEMKRQSIQED 221

RESULT 6  
US-11-097-143-39276  
Sequence 39276, Application US/11097143  
Publication No. US20050208558A1  
GENERAL INFORMATION:  
APPLICANT: Ventec, J. Craig  
APPLICANT: et al.  
TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID  
TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE  
TITLE OF INVENTION: DROSOPHILA GENES.  
FILE REFERENCE: C1000728  
CURRENT APPLICATION NUMBER: US/11/097,143  
CURRENT FILING DATE: 2005-04-04  
PRIOR APPLICATION NUMBER: 60/157,832  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: 60/160,191  
PRIOR FILING DATE: 1999-10-19  
PRIOR APPLICATION NUMBER: 60/161,932  
PRIOR FILING DATE: 1999-10-28  
PRIOR APPLICATION NUMBER: 60/164,769  
PRIOR FILING DATE: 1999-11-12  
PRIOR APPLICATION NUMBER: 60/173,383  
PRIOR FILING DATE: 1999-12-28  
PRIOR APPLICATION NUMBER: 60/175,693  
PRIOR FILING DATE: 2000-01-12  
PRIOR APPLICATION NUMBER: 60/184,831



QY 231 NPVOLETERBVRKAPAPVLSPKTGVNLNENTVSAGKDLSTSPK----- 275  
DB 274 NPVOLETERBVRKAPAPVLSPKTGVNLNENTVSAGKDLSTSPKCHVCTLPBPHCMCK 333  
QY 276 -----PSPISP 282  
DB 334 HMTDPAITLPLP 345

RESULT 8  
US-10-473-576-5  
Sequence 5, Application US/10473576  
Publication No. US20040101884A1  
GENERAL INFORMATION:  
APPLICANT: INCYTE CORPORATION  
APPLICANT: LU, DYUNG AINA M.  
APPLICANT: KAVITZU, CHANDRA S.  
APPLICANT: GANDHI, AMEENA R.  
APPLICANT: HAFALIA, APRIL J.A.  
APPLICANT: DING, LI  
APPLICANT: LU, YAN  
APPLICANT: RAMKUMAR, JAYALAXMI  
APPLICANT: SWARNAKAR, ANITA  
APPLICANT: TANG, Y. TOM  
APPLICANT: YUE, HENRY  
APPLICANT: TRAN, BAO  
APPLICANT: LEE, SOO YUEN  
APPLICANT: WARREN, BRIDGET A.  
APPLICANT: NGUYEN, DANNIEL B.  
APPLICANT: THANGAVELU, KAVITHA  
APPLICANT: YAO, MONIQUE G.  
APPLICANT: ELLIOTT, VICKI S.  
APPLICANT: BAUGHN, MARIAH R.  
APPLICANT: EMERLING, BROOKE M.  
APPLICANT: LAL, PREETI G.  
APPLICANT: GIETZEN, KIMBERLY J.  
APPLICANT: BECHTA, SHANYA D.  
APPLICANT: MARQUIS, JOSEPH P.  
APPLICANT: KABLE, AMY B.  
TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT  
FILE REFERENCE: PP-0921 USN  
CURRENT APPLICATION NUMBER: US/10/473,576  
PRIOR FILING DATE: 2003-09-29  
PRIOR APPLICATION NUMBER: PCT/US02/09809  
PRIOR FILING DATE: 2002-03-29  
PRIOR APPLICATION NUMBER: US 60/280,387  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: US 60/282,335  
PRIOR FILING DATE: 2001-04-05  
PRIOR APPLICATION NUMBER: US 60/286,663  
PRIOR FILING DATE: 2001-04-13  
PRIOR APPLICATION NUMBER: US 60/285,484  
PRIOR FILING DATE: 2001-04-19  
PRIOR APPLICATION NUMBER: US 60/350,702  
PRIOR FILING DATE: 2002-01-18  
PRIOR APPLICATION NUMBER: US 60/351,749  
PRIOR FILING DATE: 2002-01-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PERL Program  
SEQ ID NO 5  
LENGTH: 1523  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No: 1510943CD1  
US-10-473-576-5

Query Match 17.2%; Score 945.5; DB 4; Length 1523;  
Best Local Similarity 30.9%; Pctd. No. 36-42;  
Matches 258; Conservative 84; Mismatches 147; Indels 345; Gaps 16;  
QY 266 AGKDLSTSPK-----PSPISPVLGRKPNASGLLWVCKEVTIGYRGVKTITFTSM 317

DB 1010 AGAPRASPEKAEEDRRLLPGSOAPAL-----VSSQGLLEWCEVTTGYGVRITFTSM 1066  
QY 318 RNLSCATLHHRPRLIDYKSLNPDIDKNNKKAADGSAISIGSLLEPSDVLAIAD 377  
DB 1067 RNLGAFCAILHRRYPDKIDYASLDPLNIDKNNKQAADGSAISIGSLLEPSDVLAIAD 1126  
QY 378 KLTVMYTLVQIRAHFSGQELNVVQIENSSKTYKGVNTEPDNSVD-----QEKFYAEL 433  
DB 1127 KLTVMYTLVQIRAHFSGQELNVVQIENSSKTYKGVNTEPDNSVD-----QEKFYAEL 433  
QY 434 SDIKREP-ELQOPISGAVDFLSDSDSVFVNDGSGVSSSEHQTDPDHLSPSTASPYCRRT 492  
DB 1187 AEGQEPKEAADRADGAAPGAVASRNVAAGRASDQGAHA----- 1225  
QY 493 KSDTEPQSQ-----QSSGRTSGSDPDGICSTNDSTQAVLLGKKRLKAELELSPLYVS 548  
DB 1226 -----PRESRPAEVPAPGAVNGAGAG-----GGVRLRRPSV----- 1258  
QY 549 DKKKDMSPPICETDEQKLTLDIGSNLEKELNSRSLGCRSDPESP1KKTSLSPSK 608  
DB 1259 NGEPSGVP----- 1271  
QY 609 LGYSYSRDLDLAKKQASLDQTES-----DPDADRTTLNADHSSKIYQHRLLSROBELKE 664  
DB 1272 GSFSSHVRDADLLKKRSRLRNSSSFEMDDPA----- 1303  
QY 665 RARVLLEQARRDALAKGNKHTNTATPFCNRQLSDQDEERRQLREARQLIAEABSG 724  
DB 1304 ----- 1303  
QY 725 VKMSLEPYGEMAAEKLKERSKASGDENIDETIDNEBIEPGFVVGSGDELNTLNDLDT 784  
DB 1304 -----GAMGA-----AAEGAPD----- 1317  
QY 785 PQNSKLVLDKULLEVPQVANSPPSSAQAATVSSSQDKSGTEDIUTERLQKTER 844  
DB 1318 -----PSPAPGPPTA----- 1328  
QY 845 FRNPVVPKSDSTVVRKTQLQSFQYIENRPMRGQSIQDITKKNDEKAAITETQKPS 904  
DB 1329 -----DSQOPGSGSPSEBPPSPGE 1349  
QY 905 DEVLNKGFDTSQYVYVGEIALLNEBOKQIDTPAALVEKRYLIMPGRYTEEBAWQEM 964  
DB 1350 EAGLGR-FQDTSQYCAELQALFQEKQKQIDGSAAYEMQJRLBSGANKLOBEVLIQEM 1408  
QY 965 FVLVKKKALIRRMQQLSLBEKHDLERRYELNRELRLAMLAIDWQKTEAQKREQLLL 1024  
DB 1409 FVLVKKKALIRRMQQLSLBEKHDLERRYELNRELRLAMLAIDWQKTEAQKREQLLL 1468  
QY 1025 DELVALVKKKALVYRDLDQAEQAEEDHLEERTLEFQNGKNAK--KEEKCYL 1075  
DB 1469 BEIVSLVQKDELVDLDHKEKRIALBEDERLBERLGEORRRKLSRLSRRCVYL 1522

RESULT 9  
US-09-864-761-42973  
Sequence 42978, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Ranzel, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aeonica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
PRIOR FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42978
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007629.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.98
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.1
; OTHER INFORMATION: EST HUMAN HIT: BF307519.1, EVALU8 2.00e-41
; OTHER INFORMATION: SWISSPROT HIT: P32583, EVALU8 8.10e-02
; US-09-864-761-42978

Query Match      16.3%; Score 897; DB 3; Length 175;
Best Local Similarity 100.0%; Pred. No. 8.1e-41;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      352 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTTTYIYQIRAHFSGGELNVVQIEENSSKSTY 411
DB      1 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTTTYIYQIRAHFSGGELNVVQIEENSSKSTY 60
QY      412 KVGNYETDNTSSVQOEKRYAELSDIKRPELQOPISGAVNDPLSDDDSVFVNDSGVGSSES 471
DB      61 KVGNYETDNTSSVQOEKRYAELSDIKRPELQOPISGAVNDPLSDDDSVFVNDSGVGSSES 120
QY      472 EHQTDDHLSPTSPYCRRTKSDTEPKSQSSGRTSGSDDPGICSNVSTQAO 526
DB      121 EHQTDDHLSPTSPYCRRTKSDTEPKSQSSGRTSGSDDPGICSNVSTQAO 175

RESULT 10
US-10-029-386-33831
; Sequence 33831, Application US/10029386
```

```

; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hank, David R.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AROMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33831
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007098.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.61
; OTHER INFORMATION: SWISSPROT HIT: P32583, EVALU8 1.10e-01
; US-10-029-386-33831

Query Match      16.3%; Score 897; DB 4; Length 175;
Best Local Similarity 100.0%; Pred. No. 8.1e-41;
Matches 175; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      352 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTTTYIYQIRAHFSGGELNVVQIEENSSKSTY 411
DB      1 AYDGFASIGISRLLEPSDMVLLAIPDKLTWTTTYIYQIRAHFSGGELNVVQIEENSSKSTY 60
QY      412 KVGNYETDNTSSVQOEKRYAELSDIKRPELQOPISGAVNDPLSDDDSVFVNDSGVGSSES 471
DB      61 KVGNYETDNTSSVQOEKRYAELSDIKRPELQOPISGAVNDPLSDDDSVFVNDSGVGSSES 120
QY      472 EHQTDDHLSPTSPYCRRTKSDTEPKSQSSGRTSGSDDPGICSNVSTQAO 526
DB      121 EHQTDDHLSPTSPYCRRTKSDTEPKSQSSGRTSGSDDPGICSNVSTQAO 175

RESULT 11
US-11-097-143-23565
; Sequence 23565, Application US/11097143
; Publication No. US20050208558A1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT: et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION: ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; FILE REFERENCE: CU000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; PRIOR FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23565
; LENGTH: 996
; TYPE: PRT
```



Qy 698 LSDQDEERRRQRRARQOLIAEARSQVMSSELPYSGMAAEKLEKERSKASG--DENUNI 755  
Db 437 LPKQMD-----LESPKQVLE-----TKASETDYEDLNPKODDSSKGANPFSSD- 484  
Qy 756 EIDNBEIP-----EGFVVGSGDE-----LTINENDLTPEQNSKLVDLKTKLEFVQOV 806  
Db 485 EVELLKALPAQOSKQKVPVPPPPPKIGLISISINSEKPHSPPTL--SHGKKMPPTPRI 542  
Qy 807 ANSPSSAAQKAVTESSEQDMKSGTEDLTERLQ--KTERF-----RNP 848  
Db 543 SIKTQTPAKMT--HOGQKSSISSSSSEHLNSTRDRGADRGSSISLPSANGPKP 599  
Qy 849 VVFSKSTVRK-----TOLQSFQYIENRPEMKORS-----IQED-----TKKNEEK 892  
Db 600 LRASVGSPLRSEBSPTLSISITS-----PWKKRQAPLPPIQTPDPSDGFSLSDQ 654  
Qy 893 AAITETQK-----PSEDEVANKGFQTSQY----- 918  
Db 655 KALHTQLKAPNLGDSRRLPLDQSLSDSEATESNYESLSTSNADENVVYRILV 714  
Qy 919 ----- 918  
Db 715 PPTQPTENTVERSKEDQSPYVYNDPDRNVSPLGANKSTHGKRRKRGAPAVPIPRKVL 774  
Qy 919 -----VGEIALALENEQKQIDTRPALVERKL-----YLMDTG-----RNTSE 956  
Db 775 QRLPLQERHFEIIVQQLGKGVLEKMGIRCRSLDADYTDPEGAELVITNSKE 834  
Qy 957 BEAMQEWMLVNKKNALIRPMNOLSLKEKHDLERRYELLNREIRAMALIEDNQTEAQ 1016  
Db 835 VEDTILQFELVYNKELFRHQAEIMYLRQHRLEQOADIENHETRVLMQPEHNKTDSD 894  
Qy 1017 KRRQQLLDELVALVYNKDALVRLDQAEKQAEEDHEF-----RTL 1059  
Db 895 KAHSEVILNRLVKVEMNEVILSDLETRVBARDSIKNRLHYNSEREEPPAPHS 954  
Qy 1060 EONKGMKAKKEK 1072  
Db 955 DKSSKLSKKEK 967

RESULT 13  
US-10-359-012-2  
; Sequence 2, Application US/10359012  
; Publication No. US20030232419A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: KOLODKIN, Alex L.  
; APPLICANT: TERMAN, Jon R.  
; APPLICANT: MAO, Tianyi  
; APPLICANT: PASTERKAMP, Ronald J.  
; APPLICANT: YU, Hung-Hsiang  
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT  
; TITLE OF INVENTION: AND METHODS OF USING THE SAME  
; CURRENT APPLICATION NUMBER: US/10/359,012  
; CURRENT FILING DATE: 2003-02-04  
; PRIOR APPLICATION NUMBER: US 60/388,325  
; PRIOR FILING DATE: 2002-06-13  
; PRIOR APPLICATION NUMBER: US 60/384,302  
; PRIOR FILING DATE: 2002-05-30  
; PRIOR APPLICATION NUMBER: US 60/354,178  
; PRIOR FILING DATE: 2002-02-04  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 863  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-359-012-2

Query Match 9.1%; Score 498.5; DB 4; Length 863;  
Best Local Similarity 24.9%; Pred. No. 1.9e-18;

Matches 218; Conservative 137; Mismatches 367; Indels 153; Gaps 30;  
Qy 293 SLVWCKEVTNKGKVTINTTSMRNGLSICALTHHFRPDLIDYKSLNPDIDENKKA 352  
Db 7 ALLAMCRQCGYGVGEIRDLSSFRDGLACALHRRPDLIDFSLSKONVENNNLA 66  
Qy 353 YD-GFASIGISRLLEPSDMVLLAIPDKLTWTVLYQIRAHPSG----- 394  
Db 67 FEVAEKELGIPALDPPDMWSMVPDCLSTMTYVSQYNNHCSTGQAGVSPRKLAPCS 126  
Qy 395 -----QELNVQIENSSKSTYKGVNEDTNS-----VDQEKFYA 431  
Db 127 PPSVAPTPVEBEDVAQGEELSSGSLSEQGTQTPSSCAQCVHLYQRYLAGRLYHR 186  
Qy 432 ELSDLKEBPELQGIISGAVDPLSGDDSVFVNDVSGVSESHQCPDDHLSTASPYCR 491  
Db 187 HCFRCRCSSTLPL--GAYE--NGPESETPVC-----AHCA--RLGAGTSG--TR 230  
Qy 492 TKSDEPQKSGQSGRTSGSDPGICENTDSTQAVLLGKRLKAETLELSDLYVSDPK 551  
Db 231 PGFSGQKQHQOQLADADVDVPGGSSAPAGAEADGPASPEARQITKRPVQKL 290  
Qy 552 KDM-SPPF-----ICEETDEQKQLTLDIGSNLEKCL--ENSRSLSECRSDPEPIK 599  
Db 291 QELASPAGRPPTAPARKASESTTPAPPTPRPSLQENLVEQAGSSILVNGRLHELVP 350  
Qy 600 KTSLSPTSKGYSTRD-----LDA-----KKKASLQTES--DPDADRTTLNADHSSK 649  
Db 351 KPRGTPPSEGTAPAPRDPWITLVQAEPKKRPAPLPSSSGPPSDSRQVENGC--TEE 409  
Qy 650 IVQ-----HRLSROELKERARVLLBOARDALAKANKNT--NTATPFCNRQ 697  
Db 410 VAQSPFASLESKYNPFEEHEDEKEEAPASLANSPL--GHPESTPSLHPWGIT 467  
Qy 698 LSDQDEERRRQRRARQOLIAEARSQVMSSELPYSGMAAEKLEKERSKASGENDNIEI 757  
Db 468 PTSAPTKKQAPAPAPASPLALASRLSHSEPPSATPPPALSVESLSSESASQAGEL 527  
Qy 758 DTNEIPEGFVVGSGDELITNENDLTPEQNSKLVDLKTKLE-----VQPVANSFSS 812  
Db 528 LEPAPVK-----SSSEPAVHAPGTPGNFVSLSTNSLSLSEGLVEPVEQWFOA 577  
Qy 813 A--AOKAVTESSEQDMKSGTEDLTERLQKTERFPNPFVFSKSTVTRKQLOQSFQYI 869  
Db 578 SGLAPRTSSSGQPAKPCSGATPTPLVGD--RSPVPSGSS--POLVKSCK 631  
Qy 870 ENRPEMKORSIOEDTKKGNEKAITYETORKSEDEVLNKGFK-----DTSQYV-- 920  
Db 632 ENPFRKPSAPASATKATKS-----KVPAPAPGHGFLIRKTVQADQYIIBED 683  
Qy 921 --GELALENEQKQIDTRPALVERKLYLMDTGRNTEBEAMQEWMLVNKKNALIRRM 978  
Db 684 IHGMDTIERRLDLERGVLLBEKRLGLNGEG--EDMDLVDMFKLTHEKILVARE 739  
Qy 979 NQSLKEKHDLERRYLNRBELRAML--ALIEDQKTEAKRRROLDELVALVYNKRA 1036  
Db 740 SELTYFKQQLRQADVBYELKCLNKPKCD--TEBPARKVLMQELVTLIBORNA 797  
Qy 1037 LVRLDIAQKQAEDEHLERTLEQNGKAKKEE 1071  
Db 798 IINCLDEDRQREEDQMLE-----AMIKKKE 824

RESULT 14  
US-10-359-012-14  
; Sequence 14, Application US/10359012  
; Publication No. US20030232419A1  
; GENERAL INFORMATION:  
; APPLICANT: THE JOHNS HOPKINS UNIVERSITY SCHOOL OF MEDICINE  
; APPLICANT: KOLODKIN, Alex L.  
; APPLICANT: TERMAN, Jon R.  
; APPLICANT: MAO, Tianyi  
; APPLICANT: PASTERKAMP, Ronald J.

```

; APPLICANT: YU, Hung-Hsiang
; TITLE OF INVENTION: MOLECULES INTERACTING WITH CASL (MICAL) POLYNUCLEOTIDES, POLYPEPT
; FILE REFERENCE: JHU1840-3
; CURRENT APPLICATION NUMBER: US/10/359,012
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/388,325
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 60/384,302
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 60/354,178
; PRIOR FILING DATE: 2002-02-04
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 863
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-359-012-14

```

Query Match 9.1%; Score 498.5; DB 4; Length 863;

Best Local Similarity 24.9%; Pred. No. 1.9e-18;

Matches 218; Conservative 137; Mismatches 367; Indels 153; Gaps 30;

```

QY 293 SLVWCKEVTKNYRGVKTINFTTSMWNGSLFCAILHHRPDLIDYKSLNPDIKENNKA 352
DB 7 ALLAMCRQCEGYVEIRDLSSFRDGLAFCAILHHRPDLIDPDSLKDYNFENRRLA 66
QY 353 YD-GFASIGISRLBPSDWLAIAPDKLTWMTLYOIRAHFSG-----VDQEKFYA 394
DB 67 FEVAEKELGIPALDDPMDVMSVPCLSIMTYVSQYTNHFCSPQAGVSPRKGILAPCS 126
QY 395 -----OELNVQIEENSSKSTYKVGNYETDNTSS-----VDQEKFYA 431
DB 127 PPSVAPTPVEPEVDVAQGEELSSGLSEGTGQTPSSSTCAACQOHVHLYOYTLADGRLYHR 186
QY 432 ELSDLRBEPLOPISGAVDFLSQDDSVFVNDGVSSESHQTPDDHLSPTASPYCR 491
DB 187 HCFRCRCSTLLP--GAYE-NGPEEGTFVC-----AEHCA--RLGPGTRSG--TR 230
QY 492 TKSCTEPOKQSGSGRTSGDDPGICSNITDSTQAVLIGKRLKATLELSDLYVSDK 551
DB 231 PGPSQKQOHOQOLABDAKVPGGSSAPAGAADGPASPEARPOIPTYKRPVPGKL 290
QY 552 KDM-SPPF-----ICEETDEQKQLTLDIGSNTLEKEKL---ENSRLECRSDPESPIK 599
DB 291 QELASPPAGRPAPAKASSTTPAPTPRPSLSQOENLVQAGSSSLVNGRLHLPVP 350
QY 600 KTSLSFTSKLGYSYSRD---LDLA---KKKIASLRQTES--DPDADRTLLNHADHSK 649
DB 351 KPGTPEKPEEGTPAPRKDPWITLVQAEPKKPAPLPPSSPGPPSQDSRQVENGG--TEE 409
QY 650 IYO-----HLLSRQELKERARVLLLEQARDAALKAGNKNT--NTATPPCNQ 697
DB 410 VQAPSPFTASLESKPYPFEEBEDKEEAPAPSLATSPAL--GHPESTPKSLHPWYGIT 467
QY 698 LSPDOEERRROLREAROLIAEARGVMSSELPSYGEVMAAEKLEKRSKASGDENDNIBI 757
DB 468 PTPSPPTKGRPARPARASASPLALHASTLSHSEPPSPTSPALVESLSSASACTGACAE 527
QY 758 DTNEKIPGEGFVVGDELITNLENDLTPQNSKLVLDLKKCLE---VQOPVANSPPS 812
DB 528 LEPFAVPK-----SSSEPAVNAFGTPGNPVLSSTNSLASGSELVPRVEQMPOA 577
QY 813 A---AOKAVTESSEQDMKSTEDLRTBLOKTERFRNPVYVSKDSTVKTQLOQSFQY 869
DB 578 SPGLARTGSSGSPQAKPCSGATPTPLLVGD---RSPVSPGSSS---PQLOVYSSCK 631
QY 870 ENRPEKRORSIGEDTYKGINEEKAALTETORKEPSEDEVANKGPK-----DTSQYV--- 920
DB 632 ENPFRNKPSPAPASPAIKKATKS-----KVPFRPAPGHPPLIKRKQVADQYIPEED 663
QY 921 --GELALLENOKQIDTRALVYKRLRLYIMDTGRNTEEBEAMQEWFMVLVKKNALIRRM 978

```

```

DB 684 IHGEMDTERRLDALBHRGVLLSEKLGLENRSG---FDDMLVDFKLIHEKLLVRR 739
QY 979 NQSLSEKENDLRRRYLLRELRLAM--AIDEMQTEAKRREQLLDBLVALNVRDA 1036
DB 740 SELLYFKQOOLROADVVEYELRCLINKPEKWM--TEEDRARKVIMQBLVTLIEGRNA 797
QY 1037 LVRLDIAOEKQAEDEHLERTLEQNGKMAKKEE 1071
DB 798 IINVLDEDRQREBEDKMLE-----MIKKKE 824

```

RESULT 15

US-11-169-041-167

Sequence 167; Application US/11169041

Publication No. US20060019284A1

GENERAL INFORMATION:

APPLICANT: Bristol-Myers Squibb Company

TITLE OF INVENTION: IDENTIFICATION OF POLYNUCLEOTIDES FOR PREDICTING ACTIVITY OF

TITLE OF INVENTION: COMPOUNDS THAT INTERACT WITH AND/OR MODULATE PROTEIN TYROSINE

TITLE OF INVENTION: KINASES AND/OR PROTEIN TYROSINE KINASE PATHWAYS IN LUNG CANCER

FILE REFERENCE: 10001 NP

CURRENT APPLICATION NUMBER: US/11/169,041

CURRENT FILING DATE: 2005-06-28

PRIOR APPLICATION NUMBER: 60/584,405

PRIOR FILING DATE: 2004-06-30

NUMBER OF SEQ ID NOS: 527

SOFTWARE: PatentIn version 3.2

SEQ ID NO 167

LENGTH: 863

TYPE: PRT

ORGANISM: Homo sapiens

US-11-169-041-167

Query Match 9.1%; Score 498.5; DB 6; Length 863;

Best Local Similarity 24.9%; Pred. No. 1.9e-18;

Matches 218; Conservative 137; Mismatches 367; Indels 153; Gaps 30;

```

QY 293 SLVWCKEVTKNYRGVKTINFTTSMWNGSLFCAILHHRPDLIDYKSLNPDIKENNKA 352
DB 7 ALLAMCRQCEGYVEIRDLSSFRDGLAFCAILHHRPDLIDPDSLKDYNFENRRLA 66
QY 353 YD-GFASIGISRLBPSDWLAIAPDKLTWMTLYOIRAHFSG-----VDQEKFYA 394
DB 67 FEVAEKELGIPALDDPMDVMSVPCLSIMTYVSQYTNHFCSPQAGVSPRKGILAPCS 126
QY 395 -----OELNVQIEENSSKSTYKVGNYETDNTSS-----VDQEKFYA 431
DB 127 PPSVAPTPVEPEVDVAQGEELSSGLSEGTGQTPSSSTCAACQOHVHLYOYTLADGRLYHR 186
QY 432 ELSDLRBEPLOPISGAVDFLSQDDSVFVNDGVSSESHQTPDDHLSPTASPYCR 491
DB 187 HCFRCRCSTLLP--GAYE-NGPEEGTFVC-----AEHCA--RLGPGTRSG--TR 230
QY 492 TKSCTEPOKQSGSGRTSGDDPGICSNITDSTQAVLIGKRLKATLELSDLYVSDK 551
DB 231 PGPSQKQOHOQOLABDAKVPGGSSAPAGAADGPASPEARPOIPTYKRPVPGKL 290
QY 552 KDM-SPPF-----ICEETDEQKQLTLDIGSNTLEKEKL---ENSRLECRSDPESPIK 599
DB 291 QELASPPAGRPAPAKASSTTPAPTPRPSLSQOENLVQAGSSSLVNGRLHLPVP 350
QY 600 KTSLSFTSKLGYSYSRD---LDLA---KKKIASLRQTES--DPDADRTLLNHADHSK 649
DB 351 KPGTPEKPEEGTPAPRKDPWITLVQAEPKKPAPLPPSSPGPPSQDSRQVENGG--TEE 409
QY 650 IYO-----HLLSRQELKERARVLLLEQARDAALKAGNKNT--NTATPPCNQ 697
DB 410 VQAPSPFTASLESKPYPFEEBEDKEEAPAPSLATSPAL--GHPESTPKSLHPWYGIT 467
QY 698 LSPDOEERRROLREAROLIAEARGVMSSELPSYGEVMAAEKLEKRSKASGDENDNIBI 757
DB 468 PTPSPPTKGRPARPARASASPLALHASTLSHSEPPSPTSPALVESLSSASACTGACAE 527
QY 758 DTNEKIPGEGFVVGDELITNLENDLTPQNSKLVLDLKKCLE---VQOPVANSPPS 812
DB 528 LEPFAVPK-----SSSEPAVNAFGTPGNPVLSSTNSLASGSELVPRVEQMPOA 577
QY 813 A---AOKAVTESSEQDMKSTEDLRTBLOKTERFRNPVYVSKDSTVKTQLOQSFQY 869
DB 578 SPGLARTGSSGSPQAKPCSGATPTPLLVGD---RSPVSPGSSS---PQLOVYSSCK 631
QY 870 ENRPEKRORSIGEDTYKGINEEKAALTETORKEPSEDEVANKGPK-----DTSQYV--- 920
DB 632 ENPFRNKPSPAPASPAIKKATKS-----KVPFRPAPGHPPLIKRKQVADQYIPEED 663
QY 921 --GELALLENOKQIDTRALVYKRLRLYIMDTGRNTEEBEAMQEWFMVLVKKNALIRRM 978

```



Db 468 PTSSPKTKRPPAPRAPSPLALHASRLSHSEPPSAPSPALSVESLSSESASQTAGAEI 527  
QY 758 DTNEIRPGFVVGSGDELTNLENDLDTPEONSKLVDLKLKLLB-----VOPQVANSPPS 812  
Db 528 LEPNAVPK-----SSSEPVAHPGTPGNPVSLSSTNSSLASSGELVEPRVQMPQA 577  
QY 813 A--AQKAVTESSEODMKSGTEDLNTERLOKTERFRNPVVFSSKDVTKTQLOQSPSQYI 869  
Db 578 SPGLAPRTRGSSGCPAKPCSGATPTPLLVGD---RSPVSPGSSS---PQLOVKSCK 631  
QY 870 ENRPEMKRORSIQEDTKKGNEEKAITETOKKPSSEDEVINKGFK-----DTSQYV--- 920  
Db 632 ENPFMRKPSPAASPATKATKGS-----KPVRRPAPGHGFPPLIKRKVQADQYIPEED 683  
QY 921 --GELALENEQKOIDTRAAALVEKRLRYLMDTGRNTEEEEMMOEWFMVLVNGKALIRRM 978  
Db 684 IHGENDTIERRIDALEHKGVLLEKRLRGGLNEGR---EDDMLVDMFKLIHKHLLVRRB 739  
QY 979 NQLSLBEKHDLERRVEYLNRBLRAML--AIEDMQTEAQKRRBOLLDELVALVNRKDA 1036  
Db 740 SELIYVFQONLEQOQADVBYELRCLINKPEKDW--TEEDRAREKVLMOELVTLIEQRNA 797  
QY 1037 LVRLDIAQEQAEEDDEHLERTLEQNKGKMAKKEE 1071  
Db 798 IINCIDBDROREBEDKMLB-----AMTKKKE 824

Search completed: May 20, 2006, 17:30:18  
Job time : 193 secs

**THIS PAGE BLANK (USPTO)**



```

Qy 620 AKKHSAJROQESDDADRTTLINADHSHKIVQHLLSNQGBELKXRAVLLEGARDAL 679
Db 440 VELRNTL-SSGTDOSVNR-----HSHQTTY-----IKHLLOLNGKLFEBTNEARI 486
Qy 680 KAGNKHNTNTATPFCNRQLSDPOQE-ERRRQLERARQLIENARSGVXMSBLSYGENMAA 738
Db 487 ---NEYNA-LAQPLVDEAIAFKEBVLATISGELRBRTOILA-----KQAEAPT-----KR 532
Qy 739 EKLKERSQA-SGDENDINIEIDTNESEIPEGFVVGSGDELTNLENDLDTPEQNSKVLDKLK 797
Db 533 EHVEIDIEIPGLEKDSILRLVNR-----SPEIENVAAYTLQI-----ITTEIN 577
Qy 798 KLEVEQOVANSPSSAAQ-----KAVTESSBODMKSGTEJLRTERLOKTTFRAPVAFS 852
Db 578 KIFPE-EVNAFPAFAVPIIDGITLAVTSSHQ-----IP 609
Qy 853 KDSIVRKTOLOSFSOYIENRPEMKRORSIQEDTKKNEKKAITETORKPSDEVLANKGF 912
Db 610 GDAAVKQAQVE-----ENLASLDCVRR-KOPSPGSIORASIOAOL-----I 651
Qy 913 KOTSGY--VGEIALALENQOIOIFRALVLERLALYIMDTGNTSEBEMQWEMFL--- 967
Db 652 KYTTYSGATEBLV-----QARLDIKQITILACK-----TLEEBERRKARINNVYKADE 700
Qy 968 -VNKNQALIRBNQOLSLKEKHDLERRRYELNRELRAM--LAIIBMOKTEAQKRRBOLL 1024
Db 701 HNNENHNLDGKINSVD-GELBPQKQKLYE-VAEVNAKKKXKAVELTJPIYEDLBKQULHL 758
Qy 1025 -----DELVALVNKRDAIYVDLDAQ--EKQAEEDBHL 1055
Db 759 EITSPPASINIFENLIIHI---DPLVEIIDARILAAAGLEISEEEL 802

```

```

RESULT 2
US-10-505-928-100
? Sequence 100, Application US/10505928
? Publication No. US20060088532A1
? GENERAL INFORMATION:
? APPLICANT: Ludwig Institute for Cancer Research et al.
? TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
? FILE REFERENCE: 28967/39178
? CURRENT APPLICATION NUMBER: US/10/505,928
? CURRENT FILING DATE: 2004-08-27
? PRIOR APPLICATION NUMBER: US 60/363,019
? PRIOR FILING DATE: 2002-03-07
? NUMBER OF SEQ ID NOS: 866
? SOFTWARE: PatentIn 3.2
? SEQ ID NO 100
? LENGTH: 2871
? TYPE: PRT
? ORGANISM: Homo sapiens
US-10-505-928-100

```

Query Match	4.2%	Score 233	DB 6	length 2871
Best Local Similarly	18.4%	Pred. No. 1e-07		
Matches 164	Conservative 164	Mismatches 297	Indels 266	Gaps 36

QY	343	QDIEKNKKAYADVGASIGISINLBSBDMVLAIPKXLTVMYTYIYQIRAHFSGGELNVQI	402
Db	1129	EDRFDQQGNDDY---QIQAKAQCEKENINGMOKLSEBKAIKEREYI-----ERLRVLLQ	1179
QY	403	EENSSKSTY-----KVGNYETDNTSSVDOEKFAEYLSDLK---REPELQOP-----ISGA	449
Db	1180	EEGRKKEEYENELAKVRRHNYEBSNL-RANKYEIRINITYKTTIKIEISMOKEDDSGNLRNQ	1238
QY	450	VDPLS-----QDDSVFVNDSGVESESEHQTPDHLSPSTASPYCRRTKSDTEBOKSOQ	503
Db	1239	LDRLSRENBDLKDEIVRLNDSITLQATEQR-----RAEENALQOKQA--	1279
QY	504	SSGRTSGSDDPGICSNDSITQAVQLGKKRLIKATLESLDLYVSDKKKMSPPICEET	563
Db	1280	-----CGSE-----IMOKKHLEIKELMOORSBNARH-----KOS	1312

QY	564	DEOQLQTL-DLGSNLKFKLENSSLEKRSPESSIKTSLSPYKGLVYSRDLDAK	622
Db	1313	LEBAKTTIQDNKEIERLKAFQOEAKRWEMEL-----SKVNNDEETISLK	1365
QY	623	KHASICOTESPDADRITLINADHSSKIVORLLASROE-----LKERAVLLE	671
Db	1364	-----NQETENINTIKTTI-----HQLTMQKEEDTSGYRAQIDNLTREMRSLSE	1407
QY	672	QARSDALKAGNKENTNTATPFCNRQLSD--QOQBERRQLREPARQLIAEARSQVM--	727
Db	1408	EIKR--LK-----NTLTOTENLRVREDIIOQKATGSEVSORQOQLEVEILROYQMRT	1455
QY	728	SELSYGEMAAEKLEKSKASGENDNTE-----ID--TNE-----	761
Db	1460	EESVRY-----KSLDDAKTTIQDNKEIERKOLIDKETNTRKCLJEDENARLQROYDLO	1515
QY	762	-----EIPGFFVVGGBDELTNLENDLTPEONSKVLDKLK-----KLAEV	802
Db	1516	KANSATETINKLKVQE-----QELTRLAIDYERVSQBERTVADQDITRFQNSIKELQD	1566
QY	803	QOPLANSBPSSAQAAVTESS-----EODMKSGEDLTERTLQDT--TERPRNVPVSK-	853
Db	1569	QKQVIEBELNLKKTASDESCKRKLEBELLEGMRRSLKEQAIKTITNLIOOLEASIVYKR	1628
QY	854	-----DSTVRKTQ-----LOSFOYIE--NRPEMKRORSIOEDTPKKGIBEEKKA	894
Db	1629	SEDDLRQORDVLDHLEKQKQRTQSELRRLSFEVALLRQOLLQOBSYKQALIRBNHPKA	1688
QY	895	ITETQRRKSEDEVINKGFKDTSQYVGEALALENFOKI-----DTRAA	938
Db	1689	IEDRSRGSINBSKIEIERLQSLTENILTEHMLLEBELRYLRLEYDLRGRSEADSKNAT	1748
QY	939	LVEKRLAYLMDGTGNT-----EEMQMOE-----WML	967
Db	1749	LTLEBSQLOJISNNFTLELOGLINDLQRRREYLRQIEFKQKQALBASNRIOESKNQTOV	1806
QY	968	VNKINALLIRNNQUS-----LLEKENDLERREYELINBELPAMLAIEDWQKTEAQKREQL	1022
Db	1809	VQEHESILLVTKIVLEQDKARLQRLJEDELINRAKSTLEABTVK-----QRLSECEKQOION	1865
QY	1023	LLEDLVALVNRKQALVRLDAQOQABEEDHLERTLEFQNGKQAKKAEKC	1073
Db	1863	DLNMWKTQYSKEAIRIKSIBERSKSRKXNSLSSEIERLOAETKRIEBC	1913

```

RESULT 3
US-10-505-928--325
Sequence 325, Application US/10505928
Publication No. US20060086532A1
GENERAL INFORMATION:
APPLICANT: Ludwig Institute for Cancer Research et al
TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
FILE REFERENCE: 28967/39178
CURRENT APPLICATION NUMBER: US/10/505,928
CURRENT FILING DATE: 2004-08-27
PRIOR APPLICATION NUMBER: US 60/363,019
PRIOR FILING DATE: 2002-03-07
NUMBER OF SEQ ID NOS: 866
SOFTWARE: PatentIn 3.2
SEQ ID NO 325
LENGTH: 3113
TYPE: PRT
ORGANISM: Homo sapiens
US-10-505-928--325

```

Query Match	4.1%;	Score 223.5;	DB 6;	Length 3113;
Best Local Similarity	19.3%;	Pred. No. 4.9e-07;		
Matches 237;	Conservative 215;	Mismatches 486;	Indels 293;	Gaps 53;

QY 41 EGCAITDEMOSLASLVSMKQADIGNL-DDFEENED-----DDENRNVQEEKAK 89

Db 1572 OGIMKNNHIOEIEKQLLESEROEIDCLRKQYLSENEQWOKITSVTLEKSKLAEEKOTE 1631

```
QY 90 ITELINKNPLDEAKDLATVNSNPPDDPA-----ELNP 125
Db 1632 QLSLEFVARLQLOGLDSSRLGI DTEDAIQGNESCDISKEHTSETTERPHGDVHQ 1691
QY 126 FGDDPSE-----EPITETASPKRTEDSPYNS-----YNPKKQV 161
Db 1692 ICDKAOQDLMIDIEKITEGTGALKPTGSCGEOQSDTYNPEPGEDKTOGSSCISLSLPS 1751
QY 162 -POLYNPDEBEAFYTIKDS-----POSTGRKIRIPVDMSKYLYADSSKTEBEELDES 214
Db 1752 GPNALVND-----FLGNQEDIHNLQRYKETSNEMLRL-----HYIEDRDRKVE----- 1797
QY 215 NPPEPKSTPPNNILVNPVQLEET-----ERRVKKKAPAPVLSPKTGVLMNTVYASAKD 269
Db 1798 -----SLNEMKELDSKLMHQEVQJMTKIECELEBKIVELIKKE-----NSD 1840
QY 270 LSTSPKSPISPVYGRKPNASQ-----SLVWCSEVTKYNGVKTNTTTSR----- 318
Db 1841 LSEKLEYPSCDHQELQREVTSEBGLNSDLEMHADSSREDIGDNVAKVNDSEKREFLVE 1900
QY 319 NGLS-----FCALILH-----HFRPDL--IDYKSLNPQDIKENNKGAVDFASIGISRLBPS 368
Db 1901 NELSRIBEKASIEHEALYLEADLEVOTBEKLEKNENKOK----- 1943
QY 369 DMVLAIPDKLTVMT--YDYQIRAHFSGQELNVVQIENSSKSTYKGVYETDYNSSVDQ 427
Db 1944 --VIVCEBELSVLTSEBNQRLGELDTMSKTTMLDQLSBKWKKE--TQELBSHQSECLH 1999
QY 428 KFYAELSLKRPHELOPISGAVDFLSODDS-----VFVNDSS--GVGSSESEHQTPD 477
Db 2000 CIOVAEAEVKETBELQTLSSVSELLDXKTHLOKLOSLKESDQALSJLYTCELENNQIAQ 2059
QY 478 -----DHSFSTASPYGCRKTSPTREPOKQOOSGSGTSGSDPG-----ICSNMDSYQAVL 528
Db 2060 LNKKEKLVYKESSEIQAARLSSEBYE--KLNVSKALMALVKEGEPALRLSSTOEVEHOLR 2117
QY 529 LGRKRL--LKAETLESLDYSDVKKKMSPPICEBTDEQKQTLDDIGSLEKEKLENSRS 587
Db 2118 RGIKRLVRIRIADDEKQOHIABKLERE-----RENDLSK-----DKVNLREILOMSSEN 2168
QY 588 LEC-----NSDPSPIKTSLSPYSKLGYSRDLAKKXNASLRQTESDPDARTTL 641
Db 2169 QELVILDAENSKAEVETLKTQIIEEMARSLKVFELDLVTLRSEKEMLTQIOBKQOQLSEL 2228
QY 642 NHAHSSKIYOHRLISROEEL--KERARVILJEOARRAALAKAG--KENTNTATPPNCR 696
Db 2229 D-----KLLSFKSLIIEEKEQAEIQIKESKTAVENTLOQLKELINEAVALCGD 2277
QY 697 QLSDOODE-----ERRRQLREBARQLIABRSQV-----MSLPSYGEAALAKLER 744
Db 2278 QEIMKATEQSIDPPIEBEHQARNSIEKIRARLADDEKQOLCVLOQKE--SEHNDLILGR 2336
QY 745 SK-----ASGDENDNIETDNEBIPGFE-----VVGGDDELNTL 778
Db 2337 VENLERLEIARTQOEHALAEANSKGEVETILKAKIEGATQSLRGLBLDVTTIIRSEKENTL 2396
QY 779 ENDLDTBPQNKVDLKLKLLVQPOVYANPSSAAQAVTESSEQDMKSGT--EDARTE 836
Db 2397 TNELO--KEORRISLEI-----INSSPENILQKEQEKQOMKEKSSSTAMEMLTQO 2445
QY 837 RLQKTERFRNPVVSQSTVRAKTQLOSFQYIENRPEMKQRSIQ--DTKGN----- 869
Db 2446 -LKELINREV--ALAHNDQACAKQKQNLSSQVECL--ELEKQALQLOGLDEAKKNYIVQS 2500
QY 890 -----EKAATETQKPSBEDVANKGPDTSQYVVGELALLENB-----QKOIDTRA 937
Db 2501 SVNGLIOEVEDGKQLEKDEBEISRLKQIODQO--LVSKISQYEGEHQWKQNLLEIRN 2559
QY 938 ALV--EKRLATYMDTGRNTEEBEAMQEMFVLVNNKAL--IRRNQQLSLEKHDLBERRY 994
Db 2560 LTVLEBQCTQVLOSKNMSLQJTLLEVLOSYSK--NLENELELTQKADKSPFEKVKMKAKE 2617
QY 995 ELANREILRAML--AIEDWQKTEAQKRR-----EQLLT-----DELVALVNRDALVRDL 1041
```

```
Db 2618 TELQREHEHMAQKTAELQEBELSGEKNRLAGELQLLLEIRKSSKOQLKELTLENSELKKSLL 2677
QY 1042 DAOEKQAEDEBEHLEKTLTQNKKGAKKAKEK 1072
Db 2678 DOMHDOVEKEGKVAEBEIAEYQRLHEAEKK 2708

RESULT 4
US-10-523-014-4
; Sequence 4, Application US/10523014
; Publication No. US2006009401A1
; GENERAL INFORMATION:
; APPLICANT: Linn, Lih-Liang
; APPLICANT: Yannoni, Yvonne
; TITLE OF INVENTION: M2C Interacting Proteins
; FILE REFERENCE: 08702-0097-00304
; CURRENT APPLICATION NUMBER: US/10/523, 014
; PRIOR FILING DATE: 2005-02-01
; PRIOR APPLICATION NUMBER: USSN 60/400, 044
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 915
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-523-014-4

Query Match 4.1%; Score 222.5; DB 6; Length 915;
Best Local Similarity 24.5%; Pred. No. 9.5e-08;
Matches 100; Conservative 66; Mismatches 175; Indels 67; Gaps 15;

QY 6 SPNFTQDVLYKFKPLSKKVVSAALQPSLCTPIREKQADDEDMQSLASLVSKQADIGN 65
Db 533 APSSRGGCSIXMPEBPAEPLAAVAEANGABQTRVKA--PEGRSPLSABELMTIEDEGV 591
QY 66 LDDPEBNEDDERVNOBEKAAKTEL--INKNPLD--EAEKDLATVNSNPPDDPAEL 123
Db 592 LDKKLDQSTDEBEKLLR--AAURELQKQKQKQKQKERRRLOEARGRGE----- 640
QY 124 NPFGDPSREBITETASPR-----KTEDSPYNSYNPKKQVOTPOYLNPDPBP 172
Db 641 ---GRGNATETTRRHQRAADGSAVSTVKTBLVHNSNQTARFTTYE-----S 689
QY 173 AFVTIKDSPPOSTYKRNIRPYDMSKYLYADSKTEBEELDESNPFPKSTPPNNILVNP 232
Db 690 SFVARSSENGSGSTW-----MQTKTFSSSSSSSKMGSI FPR-----EDQASPRAGSLA-- 736
QY 233 VQELTERRYRKAPAPVLSPKTGVLN-----ENTVSAKGLDSTSPKSPISPV 283
Db 737 --ALEKQOAEKKEKLMQAQSLPTISASQARKAMIEKLEKGAAS--PGGRAAVORSTS 792
QY 284 LGRKPNAS-----QSLVWCSEVTKYNGVKTNTTTSWRNGLSFCALIHFRPDLIDYKSL 340
Db 793 FG--VFNANSIKQMLDWCRAKTRCYEHVDIONFSSSSSDGMARFALVHNPPPEAFDYQL 851
QY 341 NPQDIKENNKKAYDGFAS--IGISRLLEPSDMVLAIPDKLTVMYTYIQ 387
Db 852 SPQNRQNPFEVAFSSAETHADCPQLDTEDEMVRLEBDMKCVVYTIQ 899

RESULT 5
US-11-251-465-31
; Sequence 31, Application US/11251465
; Publication No. US20060094061A1
; GENERAL INFORMATION:
; APPLICANT: Brya, Reginald
; APPLICANT: Vandeghinste, Nick
; APPLICANT: Tomme, Peter
; APPLICANT: Klaassen, Hubertus
; TITLE OF INVENTION: Molecular Targets And Methods To Identify The
; TITLE OF INVENTION: Same, Useful In The Treatment Of Joint Degenerative And
```

```

; TITLE OF INVENTION: Inflammatory Diseases
;
; FILE REFERENCE: P30,172-A USA
; CURRENT APPLICATION NUMBER: US/11/251,465
; CURRENT FILING DATE: 2005-10-14
; PRIOR APPLICATION NUMBER: 60/619,384
; PRIOR FILING DATE: 2004-10-15
;
; NUMBER OF SEQ ID NOS: 880
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 31
; LENGTH: 919
; TYPE: prt
; ORGANISM: Homo sapiens
; US-11-251-465-31

```

Query Match	3.9%;	Score 213.5;	DB 7;	Length 919;
Best Local Similarity	21.7%;	Pred. No. 3.9e-07;		
Matches 160;	Conservative 107;	Mismatches 246;	Indels 225;	Gaps 34;

Qy	436	KRBELOQPIGAVDPFISQODSVVNVNSGVESHQOTPDHLSPEAS9YACRR7SD	495
Db	152	LTKRPE-----DNHTQCOIVVVEIGI-----SEROMEQHTQLMSTPEPLTUSP	198
Qy	496	TEPOKSQOSGRTSGSDDPGICISNTSDTOAQVLTAKKRLKATU-----	540
Db	199	TPASATAPTSQIGIPTSDE-----ESRPPK---SKAILBSITVSLMATPPDPKGGV	248
Qy	541	-----ELSDLYVSDKKKDNSPPTICERTDEQIKLT-----DIGSNIKEKLENSR	586
Db	249	GTGRSASIKLIAOVSPGPRD--SPAPDPSPMERVELAPILPDNFSTPTSSRTDSCVRH--	305
Qy	587	SLECSDPSESPKKTSLSPSTGYSYSRDLJLACKKJASLR-----QTESPDADR	638
Db	306	-----SPAPSPPSP--OVLAKOYS--LVAKOSVVRASRIAKKTAEPPAA--	349
Qy	639	TTLNHADSXIYQH---RLU-----SROELERARVLLEQARRDALVAGNK--	684
Db	350	-----SGNIIICHSYLERLLNVEPVKVSSEQEPEBEAEF--VAAEPEVPMKNGNS	400
Qy	685	--HN-----TWTAPEFCNRQ---LSQDODEERRROLERAROLIAEARSQV	725
Db	401	WPHNDTEIANSTPNKPPAASSPEPSPSAQOEAKTDQADGPREPPOASARKRSYQA----	456
Qy	726	KMSLEPSYGEAAAEKTL--ERSKASDENDNIETIDINEIPEGFVVGSGDELTYLENDJT	784
Db	457	--VSELEDEQHLHEDELOPFRSKT-----PSSPCPASKV-----RPLRTFLHT	498
Qy	785	PEONSKVLDTLKLLEVPQVANSPPSSAAQAV-----TESSEODM-----	826
Db	499	VQRNQM-----TPTAPSPVYKSPFIKNTPLRMDPKERQRLERLKR	542
Qy	827	-KSGTEDLTERLOKTERFPNPNVVFESHOSVTRKTQLOSFQYIENRPEMKRQ-----RS	880
Db	543	RKEBAQRLRQKVEDKRRRLBEVYKLAKKEERLKV--LQARRVEQMKKEKKQJLEQKPAQ	601
Qy	881	IQEDTKKNEKKAITETQKRPSEBZVANKPFIQTSQYVVEELALMBEQOQIDTRALV	940
Db	602	IDETKEKAEERLLEBKAKKK-----AAAKMEVEERKROE	638
Qy	941	E--RLRLIMDTGANTTEEBAMQMBFVLVNNK---NALIRPNQOLSLJEKHLDER--Y	994
Db	639	EDARRLRMLQO-----BEERRHDE--LLOKKKEBEOERLKAABAKRLAEOQROERREQ	691
Qy	995	ELINRELAMLAIEDWQKTEAQKREOULLDELVALVNRDALVRDLDQAEQKAEEDEN	1054
Db	692	ERREQERREQERRBOERREQERRREQRLQAEQRRRBEGERLQAEERLEQERKALRLQEOQ	751
Qy	1055	LERTLQONKGMAKKEEK 1072	
Db	752	LQRELEBK-----KKEEQ 765	

```

; Sequence 1, Application US/105255621
; Publication No. US20060100418A1
; GENERAL INFORMATION:
; APPLICANT: Kiyosue, Yuko
; APPLICANT: Sasaki, Hiroyuki
; APPLICANT: Tsukita, Shoichiro
; APPLICANT: Eisai Co., Ltd.
; TITLE OF INVENTION: CULTURED XENOPUS LAEVIS CELL LINES
; TITLE OF INVENTION: EXPRESSING MUTANT ADENOMATOUS POLYPOSIS COLI GENE
; FILE REFERENCE: 082368-002400US
; CURRENT APPLICATION NUMBER: US/10/525,621
; CURRENT FILING DATE: 2005-02-22
; PRIOR APPLICATION NUMBER: PCT/JP03/10434
; PRIOR FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: JP 2002-241487
; PRIOR FILING DATE: 2002-08-22
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2829
; TYPE: prt
; ORGANISM: Xenopus laevis
US-10-525-621-1

```

Query Match	3.7%;	Score 202;	DB 6;	Length 2829;
Best Local Similarity	18.1%;	Pred. No. 1.2e-05;		
Matches 225;	Conservative 150;	Mismatches 433;	Indels 438;	Gaps 52;

```

QY 66 LDDPEONEDDENRVAOEEAAKITELINKNPLDPA-----EKOLAVNSNPPDDP 119
Db 1134 VDDDDDDTTTFSEKYSSEEQOEDETERONKNTIKAYASEHHGEQPTDYSRKSTVPS 1193
QY 120 AAEIAMP-----GDPSEBEPTETASPRKTEDSPFNNSYNPFKEVQTPQYLANPDEP 171
Db 1194 SAQSEPFYSNNSKQPKKQYSSNSNTPTSPNSNRQQLHPNSAQSHPGLNRPQIP 1253
QY 172 EAFPTIDSPPOSTKRNIR-----PVMSKLYADSSKTEBEELD-----ESN 215
Db 1254 -----NKPPTSINOETIQYCEVDPTPFSSRGSLSLSADEIEGGRNSRQESN 1305
QY 216 ---PFYEPK---STPPPNLVNVPQOELETERVRK-----K 245
Db 1306 NTLQITPEKELSAVSKGAVNERSSVHHRTTNKRLQTSNISPSDSRHKSVFSGAK 1365
QY 246 AP-----APP-----VLSPTGYLVNENTYS-- 265
Db 1366 SPSPSKGAQTPKSPPEHYVQETPLMFSCRKSGSSIDSPESHISIASVASEHMIISGII 1425
QY 266 AGKULSTSP-----KPSPIPSPLGKRPNAGSLVWCAGEVTKNGYGVITNETTS 316
Db 1426 SPSPULPSPBQOTMPSPRSKTPPPPPQYQAKDGS-----KPIYDBERGKYAK--TA 1475
QY 317 WRNGLSFCALH-----HFRPDLIDYKSLNPQIDKNNKAYDGF--ASIGISRLLEP 367
Db 1476 VHSXIQGVUQLQZADTLHPA-----TSTPDGSCASLSALSIDEP 1518
QY 368 ---SDMWLLAIPKLTMTYLYQIRAHFSGGELNVQIE---ENSSKTYKXGNVETQF 420
Db 1519 YIQKDVOLKTMPPVL-----ENDQSNKKEPEKEFLDNKAKKEDKRSBQEKDM 1565
QY 421 NSSVDQEKFYAE---ISDLKREP-----ELQOPISGA----- 449
Db 1566 LDDTDDIDILBEGIISAMPKRKSRKNKKYVQPTLPKPPPPVARKPSQLPVYLLSSQNR 1625
QY 450 -----VDPLSQDD-----SVFVNDSGVGESESHQTPDDHLSPSTASP 487
Db 1626 LQTKHNFHSHDMPREVVCVEGNPIINFSTATSLSDLTIESPSE-----PTNDQF 1676
QY 488 YCRRTKSDTEPOKSGQSGRTSGSDPGICSNSTDSQAOVLGKGLKKAETLESLDLY 547
Db 1677 NTUSSLSTLEKRTQITPEGR-----STDDTDAKSPPLPTTVLDEDKABEGDILA 1725
QY 548 S-----DKKQMSPPICEETDEQAKQLTIDGNSNLEKEKLENSRSL- CRSDPESPik 599

```

Db 1726 ECHSAMPKSKSHKRYVKIMDQ-----INHSTAATSSGNSRSMQEBTDKMKTSFVK 1778  
Qy 600 -----KTSLSPTSLGYSYSDRLDACKKHAASLQTSDDPADRTTLNH 643  
Db 1779 PMSOIGEKERLKNTELKNLNNSENQCDPRK--PSSKKPSKVAANKIPNNERTYGF 1835  
Qy 644 ADHSKIVQHR-----LLSROEBL-----KERAVLLEOARRDALKA 681  
Db 1836 APDS-----PHYTPTEGTPYCSRYDSSLSDDFEDDDIDLSKRYALKEKKTOTDQV 1891  
Qy 682 GNKNNTATATPCNQLSDQODEERRQL--REARQOLIAEARSQVKNSELPYSGEMAA 739  
Db 1892 KYKHNRAINPM-----GKDDQTPKSLGGDDQKALV-----QKPTSPSSAK 1935  
Qy 740 KUKERSKASGDBENDIEIDTNEBIEPGFVVGGBELTLENDLDTPEONSKLVLDKKL 799  
Db 1936 GTODRGATDEKMEFAL--ENTPVCF--SRNSSLSTL-SDIDQENNNKETEPLK---- 1985  
Qy 800 LEVQPOVANSPPSAOKAVTSSSEBDMKSGTEDLTERLOKTERFRAPVVFSDKSTVRK 859  
Db 1986 -----QCTSETQGLRRPQTSGYAPKSFHYED-----TPVCFSPRNSLSS 2026  
Qy 860 TQLQSPSQYIENRPEMKQRSIQEDTKKNEKKAITEORKEPSEDEVLNKGFQDTSQYV 919  
Db 2027 LSTDSBDLL-----QECISSAMPK-----KRPKS--KIKNEVGKSRNSV 2065  
Qy 920 VGEIALLNEBOKIDTRALAYEKLRYLMDTGRNTEEBAMQ-----EWFMLVNNKQAL 974  
Db 2066 GGIILA--EERPLTLDLRDI-----QSPDSNAPSPOSENPDWKAIOEGANSI 2110  
Qy 975 IRMNQSLKREKNDLERYELLNRELAMLAIEBQMTHEAQKREQLLDELVAL--V 1031  
Db 2111 VSRLLHQA-----AAGSLSRSGSSDSISLSKSCI 2141  
Qy 1032 NKRALVYDLDAOEKOEEDHELEKONKG--KMAKKEKCVLQ 1076  
Db 2142 SLGSPFLTLDXEK-----TITSNKGPILLPAKSALE 2176  
RESULT 7  
US-10-505-928-357  
; Sequence 357, Application US/10505928  
; Publication No. US20060088532A1  
; GENERAL INFORMATION:  
; APPLICANT: Ludwig Institute for Cancer Research et al.  
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES  
; FILE REFERENCE: 28967/39178  
; CURRENT APPLICATION NUMBER: US/10/505,928  
; CURRENT FILING DATE: 2004-08-27  
; PRIOR APPLICATION NUMBER: US 60/363,019  
; PRIOR FILING DATE: 2002-03-07  
; NUMBER OF SEQ ID NOS: 866  
; SOFTWARE: PatentIn 3.2  
; SEQ ID NO 357  
; LENGTH: 3256  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-505-928-357  
Query Match 3.4%; Score 185; DB 6; Length 3256;  
Best Local Similarity 19.1%; Pred. No. 0.0002;  
Matches 231; Conservative 163; Mismatches 470; Indels 344; Gaps 52;  
Qy 2 KOVASPMPT---QTDVK-LKFKPLSKVVASALQPSLSCIFLRBQKATDEDMOSIASIVS 57  
Db 1172 KAMLTPKRAGDEKDIKAFMGTPVQKLDLAGTLPS-----KROLQTKKKAQALDLAG 1226  
Qy 58 MKQ-----ADIGNLDDFEEDNEDD-----ENRVQOEKAAT--ELINKL 97  
Db 1227 FKELFQTPGHTBEVLVAAGKTKTKIPCDSPQSDPVDTPSTKORPKRSIKKADVEGLACR 1286  
Qy 98 NFLDEAEKDLATVANSNPDPDDAALANFPGDPDSEPTT--TASPRTEDSFVNNSNP 155

Db 1287 NLMPAGKAMHTPKPSVGEEDIIIF--VGFPVQKLDLTENLTGSKRPQ-----TP 1336  
Qy 156 FKELQTPQYLNPPPE-----PEAFVTK-----DSPQ-----STRKNIRPV- 193  
Db 1337 KEKQALDELTLGFYELFQTPGHTBEAVALAGKTTMPCSSPSEADPTPTSTRQPKPTLE 1396  
Qy 194 --DMSKYLVA-----DSS--KTEEBELDSNFYEPKSTPP 225  
Db 1397 KRDVQKLSALKULTQTSGETTHTDKVPGGSDKINAFRETAQKLDPAASVYTSKRIK 1456  
Qy 226 PNNLVNPOBL-----ETERRVKRAKAPAPVLSPKTG----- 257  
Db 1457 TKERAOPLDLAGKELFQTPVCTDKPTTHEKTKIKACRSQPDVPDPTSSKPSKSLR 1516  
Qy 258 -----VINENTVAGKDLSTSPSPSPSPSVYLRKP--NAOSLWVCKEYTKATYG 307  
Db 1517 KVDVEEBFALRKPTPAGKAMHT--PKPA-----VSGEKNTYAFMGTPVQKLDLTENLTG 1570  
Qy 308 VK-----ITNF-----TTSWRGLSPCALIHPRPDLIDYKSLNP 342  
Db 1571 SKRLQTPKEKAQALDELLAGKELFQTRGHTBSMTMDTKAKVACSSQPLD---DKNP 1626  
Qy 343 QDIKNNKKAAYDGPASIGISRLLEPSDMVLAIPDKLTVMTYLYQIRAFPSGOELNV--- 399  
Db 1627 ASSRRLKTS--LGKVGVK-----ELIAY-GKLT-----QTSGETTHHTE 1665  
Qy 400 VOIEENSKSTYKGTNTETDNNSSVDQEKYAEISDLKREBELQOPISGAVIDLSQDDSV 459  
Db 1666 PTCGDKSMKAFMBEPKXOILDSAAISLTGSK--RQLRTPKGSKEVEDLAGPTE-----L 1716  
Qy 460 FVNSGVGESESEHQ-----PDDHLSPTASPYCRT--KSDTEPQ-----KSQ 503  
Db 1717 FQTPSHTESMTNEKTKTVSRASQPLVDPTYSKQPKSLKADTBESFLAPRKQTP 1776  
Qy 504 SSGRTSGSDPGICSNPTDST-----QAQVLGKKRL-----KATLE----- 541  
Db 1777 SAGKAMHTPKRAVAGEEKDINTFLGTPVQKLDQPNLNGSNRRLQTRKEKAQALBELGFR 1836  
Qy 542 -----LSDLVSDKKKMBPPTCEBTDEQKQTLDIGNLBKCKLNSRSLGCRSDPE 595  
Db 1837 ELFQTPCTDNPDADEK--TTKKILCKSPQS-----DPADTPNTKORPKRSLKAVDE 1888  
Qy 596 SPIKKTSLP-----TSKLGYSYRD-----LDLAKKHAASLQTSDDPAD 637  
Db 1889 EFLAFKLTTPSAGKAMHTPKRAVAGEEKDINTFVGTPEVKLDLGNLPGSKRRPQTPYEKA 1948  
Qy 638 RT-----TLNHADHS--SKIVQRLLSROEB-----LKERAVLLEOAR- 674  
Db 1949 KALEDLAGFKELFQTPGHTBSMTDDKITVYSCSPQDPVKTPTSSKORLAKISLGKGV 2008  
Qy 675 RDAALKAGKNNTATATPCNROLSDQODEERRQLREARQOLIAEARSQVKNSELPYSG 734  
Db 2009 KEVLPVQKLTQTSGETTHQTHRETAG--DGKSIKAFKESAKQMDPANYGGMERMPRT 2066  
Qy 735 EMAEKLKERS--KASGDBENDIEIDTNEBIEPGFVVGGBELTNL-----ENDLTPRE 786  
Db 2067 KEKQSLDELLAGFELFQTPDHTEESTTD-----DKTKYICKSPPESTMPTPT 2115  
Qy 787 QNSKLVLDKLELVQPOVANSPPSAOKAVTSSSEBDMKSGTEDLTERLOKTERFR 846  
Db 2116 STRRPTPLGKNDIV-----ELSAKQLOTHTTDDKVPGBEDGGINVFRATAQKL 2168  
Qy 847 NPVVFSKDSVTRKTQLQSPQYIENRPEMKR--ORSIQEDTKKNEKKAAL----- 895  
Db 2169 DPAASTVSGRKQPTPKGAQPLDLAGLKELFQTPVCTDKPTTHEKTKIKACRSQPD 2228  
Qy 896 --TETQKPSBEDVLNKGFKDQTSYVVGEIALLNEBOKIDTRALAYEKLRYLMDGRN 953  
Db 2229 VGTPTIFKPSKSLRK-----ADVEEELIARKRTPSVGKAMDTPKPAQGD 2275  
Qy 954 TEEBEAM 961

```

Db      2276 EKDKAKEM 2283

RESULT 8
US-10-505-928-213
; Sequence 213, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 213
; LENGTH: 1120
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-505-928-213

Query Match      3.4%; Score 184; DB 6; Length 1120;
Best Local Similarity 18.9%; Pred. No. 4.9e-05;
Matches 232; Conservative 186; Mismatches 444; Indels 364; Gaps 56;

Qy      28 AALQFSCTFLRSGKATDEDMQSLASVSMKQADIGN-----LDDPEEDNEDDDE 78
Db      47 SSLTPSLCKKGLGLOG-----SNSSSPVDFVNNKRTDLSSEHSHSKWLETCCQHSDEQPL 102
Qy      79 NRVAQEKAKITR-----LINKLNFLEBAKD-LATVNS-----N 113
Db      103 DPLPQISTRTKTSBAVDPLGNVWVKTIVLPSPGGQOMIFAKLDTMAETNSISLNG 162
Qy      114 PF--DDPDAELNP-FGD-----PDSEBITETASPRKTEDSPFNNSYNPFKEVQTPOY 164
Db      163 PLRTDLDVREEVACMGDRFSEVAASEKPIFO-----BESPSH 200
Qy      165 LNPDEBEAVTITKDSPOSTKKNIRPVDMSKTLVADSSKTEBEEL--DESNPFYEPKS 222
Db      201 L-----LEESPPNCSQ-----LHCSKSLSRTEAVREEDLVPSBSNAFL--PSS 244
Qy      223 T---PPPN-----NLVNPVOELTERRYKRAKAPAPVLSPTGVLNENTVSAGKD-L 270
Db      245 VMLSPETALAADPRVNVHVDPEEIVVHGAMERENKAPPTHKPSSETEDDALVSSVEDIL 304
Qy      271 STSPKSPPI-----PSP-----VLGRKPNASOSLLVWCKEVTNRYGVKLTNFTTS 316
Db      305 STCLTPVLVEMESQEGAPVAVEDVGRILGSDTBSGMSPLAWLB-----KGVNTSVMLEN 358
Qy      317 WRNGLSF-----CALLHNRPLDIYKSLNPDQIKNNKAKADGPAISIG 360
Db      359 LRQSLSLSPMLRDAIGTTPFSTCSVGTFPTPSA-----POKSTNTSQT---GLVG 407
Qy      361 ISRLBESDWMVLAIPDKLTVMTYLYQIRAHFSGQELNVVQIBNSSKSYKVGNYETDT 420
Db      408 TKHSTSTEBQLGCRPDLTALS-----RH-----DLEBNL-----L 439
Qy      421 NSSVDEKFYAEISDLKREBELQOPISGAVDFLSQDDSVFVNDSGVSESEHQTPDDL 480
Db      440 SSVLYVFFLSRQLRDMKSGQLAVPHR-----ETQDSSTQTDTSHGSIYKNGLOHKLKSH- 491
Qy      481 SPSTASYCYKRTASDTPEQKSSQSSGRTSGSDPDGICSTNDSYQAOVLGKRLKLAETL 540
Db      492 -----EWGQALQQA-----RNVWQSWVLISKELI---SL 518
Qy      541 ELSDLVYSDKKKM-----SPPTICEETD---EOKLOTLDIGSNLJBKLENSRSL 569
Db      519 HLLSLHLEBKTYNQSRRRAETLYVCCCFDLKLRKQLQSL-----KAEBEBAHREMA 574
Qy      590 CR--SDPESPIKKTSLSPSTKLGYSYRD-----DLAKKGHSLRQT 630

```

```

Db      575 LRGDAAEIVLEAFCAHSAQRISQLEQDLASMEFRGLKDAQOTQVLGHAKQBEIVQT 634
Qy      631 ESDPADRTLTINADSHSKVQH-----RLTSROBELKERARVLEQA--RRDAALKGNK 684
Db      635 VS-----LTSTLQODWRSQMDYTTWTALTSRRQRLTEKLTVVSQQLQORDAIE--EK 687
Qy      685 HNTVATPFCNRL-----SDQODEERRRQLREBARQLIAEARSQVMSLPSYGEAA 738
Db      688 QEVSRLVEQVSQALEBKQGTQLELENRLADLRQLOIILANMSQLELSQHTHCA 747
Qy      739 EKLERSKASGDENDNIETNEIPEGFVVG-----GDELTNLENDLTPE 786
Db      748 QDLAMQDELLCOLTOSNEEQAOQVKEEMALKMQAELQQQCAVLAKEVADLKETLEFAD 807
Qy      787 QNSKL-----VDLKKLLEVPQVANSPSAAQAVTESSEQDKSGTEQLRTERLQ 839
Db      808 QENQVAHLBVGVECOQKLTLEVLRL--RLQCEMLDTVEENTLAKLASTIANQEDDL 865
Qy      840 KTTFRFRNPVPSKDSIVRTQLOSPSQYIENRPEMKRQRIQEDTYKNEBRAALITETQ 899
Db      866 KTRQ-----YQKUGLLEQLOSLTLPIQTK-----LKEKTEQETILLSTA 906
Qy      900 RKPSEDEVLNKGFQDYSQYVVGELALENQ-----KQIDTR--AALVEKRLRY 946
Db      907 CPPTQEHPL-----PNDRTFGISITLTAADBEPESTVPPLGSDKSAFTRVASVMSIQ--- 959
Qy      947 LMDTGNTBEEBAQMEFWLVNKNALIRRMQSLILEKHEHDERREYELNR---ELRA 1003
Db      960 -----PATTGMEBSLMEASTMTTELQSL-----CSLQDSK--EBAIRTLQKICELQA 1007
Qy      1004 MLA--EDMQTEAQKRRE-----QLLL-----DELVALVNRDALVRLDQAQEK--QA 1048
Db      1008 RLQAQEBQHQVEYKAKAEADIEKLNQALCLRYKKEKLEQVYIQNEKILBQIDISGELISL 1067
Qy      1049 EEEDEHLERTLEONKGMKAKKEKCVIQ 1076
Db      1068 REEVTHLTRSL-----RRAETETK-VIQ 1089

```

```

RESULT 9
US-10-505-928-799
; Sequence 799, Application US/10505928
; Publication No. US20060088532A1
; GENERAL INFORMATION:
; APPLICANT: Ludwig Institute for Cancer Research et al.
; TITLE OF INVENTION: LYMPHATIC ENDOTHELIAL GENES
; FILE REFERENCE: 28967/39178
; CURRENT APPLICATION NUMBER: US/10/505,928
; PRIOR FILING DATE: 2004-08-27
; PRIOR APPLICATION NUMBER: US 60/363,019
; PRIOR FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: PatentIn 3.2
; SEQ ID NO 799
; LENGTH: 1384
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-505-928-799

Query Match      3.3%; Score 182; DB 6; Length 1384;
Best Local Similarity 21.1%; Pred. No. 9.1e-05;
Matches 206; Conservative 115; Mismatches 346; Indels 310; Gaps 42;

Qy      162 PLYLNPDEBEAVTITKDSPOSTKKNIRPVDMSKTLVADSSKTEBEELDESNPFYEPK 221
Db      161 PYPVPVPAESESILVNGNHTPQATTRGPS--ACASHSSLVSSIEKLOQIMSLVLEBEG 217
Qy      222 STPEPPNVLNPVOELTERRYKRAKAPAPVLSPTGVLNENTVSAGKD--STSPKSPPI 280
Db      218 AAGKRAAATSPSPMANGRY-----LSEPT---SPGANSVSSSTVNTSPASPSLS 266
Qy      281 SPV-----LGRKPNASOSLLVWCKEVTNRYGVKLTNFTTSWRNGLSFCALIHHR 330

```



```

Db 267 SPASSGASHSPGQBEPSPVPLVPAR-----SSSYHAL----- 303
Qy 331 RPDLDYKSLNPQDIKENNKAYOGFASIGISRLLEPSD-WLLAIPKLTMTTYVOIR 389
Db 304 QPPOSREGASERFSLRKGEHPSPGRLGLTDSPPAATVLEAARAPESPL----- 359
Qy 390 AHFSGOBLVVOIEENSSKTYKGVNYETDINSVDQKFAVLESLDKRBPLOOPISGA 449
Db 360 ---GGQLPVVAI---SLSEY-----PAGGA 378
Qy 450 VDFLSQDDSVFVNDGVBESSEHQTP-----DHLSPSTASPYCARTSDTEPQ 499
Db 379 ---LSQPTSI-----PSPKQFPVPAPRNKIGTLQRPSPFRRPGSRRVLT-TSP- 427
Qy 500 KSQSSGRTSGSDPDGICSNDSQAVLGGKRLKAELELSLVYSDKKKMSPPPI 559
Db 428 -SRQLVGTTP-SD-GLATRTLQPPESPRLRRGI--DSMRPLPLSPSLRRRLS-PLP 480
Qy 560 CEETDEQQLQTLIGSNLEKEKLENSRSLBCGSDPESPIK-----KTSLSPTSLG 610
Db 481 TRITPDK-----LNREVASPRPRMAAHGASFEDESLTIGAGKRTSPSPFLG 531
Qy 611 YVSRRDLDAKKKASLQTESDPDARTTNHMDH-SSKIVQHRLSROBELKERAVL 669
Db 532 ESLA-----PHKGSFSGRLSPAYSGSLTGASPCQSPCYQRLSSGD-----LKVP 577
Qy 670 LEQARDAALKAGKNTNTATPFCNRQ-----LSQODEERRRQUREAROLIAE--AR 722
Db 578 VTRERKNSITISD--MEDDLLEYHRQROERLREQEMERLQRLFTILMCAEVSAD 635
Qy 723 SGVMSLEPSYGEAAEKLKERKASGDENDNIEIDTNEIPEGFVVGSGELTNLENDL 782
Db 636 GGPAGELPSIGETA-----ALALG----- 657
Qy 783 DTPBNSKLVDLKQKLEVOQVANSBSSAQAQVTESSBQDKSGTEJRTERLOKTT 842
Db 658 -----RRSRGLAGASGSSSE-----PGVATQGLWMSB 686
Qy 843 ERFPRPVFSDSTYKRTQLOSPSOYIENRPMKQRIQDITKKGNEKALITETOKP 902
Db 687 E-----RSDENLK-----EBSSTESTOGE 707
Qy 903 SEDEVANKPQDSQYVVGELALENEKOIOTRALVYKRLR-----YLMOTGNTEBEE 958
Db 708 HEDAPSTK-----LOGEVLALBEBDAQVGHVEQIKVAKELBOQLQESANBAEKER 759
Qy 959 AMQ-----EWFVNLKKNALIRRNQOISLE---KEHDLERYEELINRELRAMLAIED 1009
Db 760 ALLGERAERALLQKEQKAVDQLEKLVLETGIOKRDKEA--EALFETKLPFBDE- 816
Qy 1010 WQKTBQGRREQ-----LLLD--ELVALVNRDALVRLDQO---EKQABEDENL 1055
Db 817 FOQJERESRVEEBELAGQGLRSKXABLRSIARKKERTALILDSQAGIRAOAVQESRL 876
Qy 1056 ERTLEOKKGNKAKKEK 1072
Db 877 ARDNASLQQLQKEK 893

```

```

; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 50612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; PRIOR FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2587
; LENGTH: 1674
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-511-937-2587

Query Match 3.3%; Score 180; DB 6; Length 1674;
Best Local Similarity 19.9%; Pred. No. 0.00016;
Matches 206; Conservative 133; Mismatches 322; Indels 376; Gaps 46;

Qy 220 EKSTPPNNLVNPVQBLETERR-----YKKAAPAPVLT 252
Db 53 PTALEPSPBKVDQLPRSSERRSGCGTQPPARSAVAAGAARAGAPERSPLGRRV 112
Qy 253 SPK-----TGLNENTVSAGKDLSPSPSPISPVYLGRKNASQ----- 292
Db 113 SPRLCSGEGQVAVAGIAGKRGRRGD-GSRAPGRREMFELHKKDFVRQKPPADLRPD 171
Qy 293 SLVWCKEATKNGVAVITNFTTSWRGLSP-CALHHPRDLIDYSLNPQDIKENNK 351
Db 172 BEVFYCKVTEIFHNYDFFERTILCNLWSCAVTG--RP-----GLTYQELSESK 223
Qy 352 AYDFPASI-----GISRLBPSDWVLAIPDKLTVMYTYLQIRAFSGQE 396
Db 224 ARQNLQFPBPLIPLVLYLSLTRSRHLICDDIPAVVDORVVEEVTYIRANGARLQ 283
Qy 397 LNVQIEENSSKTYKGVNYETDINSVDQKFAVLESLDKRBPLOQPSGAVDFLSQD 456
Db 284 CTILEVLPFHONGFANGHV--NSVDGE----- 309
Qy 457 DSVFVNDGVBESSEHQTPDHLSPSTASPYCARTSDTEPQKQSSGRTSGSDDPGI 516
Db 310 -TITISDS-----DDBSTQSCSQNGKKQDAIDP-- 337
Qy 517 CSNDSQAVLGGKRLKAELELSLVYSDKKKMSPPICEETDEQQLQTLIGSN 576
Db 338 -----LLEFKV-----VQPTKEIHSAIYKATQISRRKLT----- 368
Qy 577 LEKEKLENSRSLBCRDPES--PIKTSLS-----PTSKLGYSYSRD 616
Db 369 FSRDKLTLFKQHC--SPQGVYIKKASISLYKIAQDPSYFPDDPPFTFSPANRR 426
Qy 617 LDLAKKHASLQTESDPDARTTNHMDHSSKTVQNR-LLSRQELK-----EAAVYLE 671
Db 427 GRPEKRTHIS---QEDNVANKQTL--ASYRSKATYKRDKLKQENKSLAFERAKUKRE 480
Qy 672 QARRDALKAGKNTNTATPFCNRQLODDEERRRQUREAROLIAEARSQV--MSB 729
Db 481 KA--DLEBAKK-----EKEDKKK--REELKIYEBERLKKGEKER 519
Qy 730 LPSYGEAAEKLKERS-----KASGDENDNIEIDTNEIPEGFVVG-----GDEL 775
Db 520 LKVEREKERKLEBKRYEYLKQMSKPRDEMCDDLKELPEPTPVKTRLPEIPFDAL 579
Qy 776 TNL-----NDLTPBQNSKLVDLKQKLLBQVQ 805
Db 580 MVLEFLNAFGLFDLQDEPDPGVTLVLEBALVNDSEGP-----LCLELFFLTALFOA 634
Qy 806 VANSPPSAQKAVTESSEQ-----DMKSGT--EDJRTB-----RLQK- 840
Db 635 IABEEBVAEQULTADTKGCSLKLDDSCITSEIRLHILASGADVTSANAKYRYQKR 694

```

```

; RESULT 10
; US-10-511-937-2587
; Sequence 2587, Application US/10511937
; Publication No. US2006008836A1
; GENERAL INFORMATION:
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.
; APPLICANT: Wohlgenuth, Jay
; APPLICANT: Fry, Kirk
; APPLICANT: Woodward, Robert
; APPLICANT: Ly, Ngoc
; APPLICANT: Prentice, James
; APPLICANT: Morris, Macdonald
; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING

```

QY 841 -----RFRPVVPSKDSYV-----RTQLOQSPSOYIEN 871  
 Db 695 GGFPAITDADCEMLRLSNPSLVKGLSSYVDLTGEEKKILHALCGULTIVSTRDIED 754  
 QY 872 RPEMKRQ-----RSIQ-EDTKGNEEKA-----ITETQKPSDEVLN 909  
 Db 755 YVDILRAQKQEFRELKAEQHKEREBAARIRKREKEKLEQEQMKKEKQELKEDBOBN 814  
 QY 910 KGFPDTSQYVVGELALLENQKQIDTBAALVERKRLVLMQGRTEREEMMQEWPVLVN 969  
 Db 815 -----STADISIG-----BEEREDDTISIESKDTQOKEL-DQDMTEDEDDPGS-----H 858  
 QY 970 KKNALIRRNQOLSLKEKHDLERRRYELNRELAMLAIEDWQKTEAQKRRROLIDEL- 1027  
 Db 859 KRG---RGRKGQNGFPEFTROBOINCVRRL-LTDEBEALQKHQRKEKELLEKIOS 913  
 QY 1028 -VALVN---KRDALVR 1039  
 Db 914 AIACTNIFPLGRDMYR 930

RESULT 11  
 US-11-118-524-2  
 ; Sequence 2, Application US/11118524  
 ; Publication No. US20060088847A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: CU WEI  
 ; TITLE OF INVENTION: ARF-BP1 AS MEDIATOR OF P53-DEPENDENT AND INDEPENDENT TUMOR  
 ; FILE REFERENCE: 19240-497US2  
 ; CURRENT APPLICATION NUMBER: US/11/118,524  
 ; CURRENT FILING DATE: 2005-04-29  
 ; PRIOR APPLICATION NUMBER: 60/610,506  
 ; PRIOR FILING DATE: 2004-09-15  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 2  
 ; LENGTH: 4373  
 ; TYPE: PRF  
 ; ORGANISM: homo sapiens  
 ; US-11-118-524-2

Query Match 3.0%; Score 167; DB 7; Length 4373;  
 Best Local Similarity 18.2%; Pred. No. 0.005;  
 Matches 206; Conservative 164; Mismatches 450; Indels 310; Gaps 41;

QY 45 TDEDMOSLASVSNKQADIGNLDDFEEDNEDDENRVQEEKAKITELINKLFLD-EA 103  
 Db 2275 SEQDAGASQDSSSNQDDPGEBAVEQBEDHD---VTQTEVADG-----DIMGGA 2323  
 QY 104 EKDLATVNSNP-----PDDPAALNPNFGDDSEEPITETASPKRTEDSFY 149  
 Db 2324 ETDSVVIAGQPEVLSQEQMVENLELLELLELDGSGSNSTIIVRSQEDSQQEVLV 2383  
 QY 150 NNSVNFPEKQVTPQ-----YLPPEDEPAFTYIKDSPQSTKRNIRPVMSKYLA 201  
 Db 2384 DEAPSNLSQASTIQANREDSMTLIDPDEBEH--TQBEDSSGSGHEDD----- 2429  
 QY 202 DS-SKTEEBELDSNPFEKPSPPPNLVNVOELTEERRVYRKADAPVLSPKTG--- 257  
 Db 2430 DSDQDEEBEEDDEDDQDEGEGBDDDDGSEMELDDYDPMNASPLVRFERFRED 2489  
 QY 258 ----VLNENTVSAKDLSTSPKSPISPVYGRKPNASQSLVWCKEYTKRYGVKITNFT 314  
 Db 2490 DLIIEDPNMSSATDIPSPGNFTYHPLVWRHADHSLTLAGSGSSSTRLTQIGISQRT 2549  
 QY 315 TSWRNGLSFCALIMHF-----RPDLIDYKSLNPQ---DIKENNKAYADGPAISGTSRL 365  
 Db 2550 LRQULTANTGHTIVHVGKRNQPNPPLILQRLLEBSAADLLOJSSSL--PLQSRGRARL 2607  
 QY 366 EPESDMV-----LLAIPDKLTWTYLYLQIRAHFSGQELN 398

Db 2608 VGNDVHIIARSDELDDFFHDQSTATSQAGTILSIPALTREWT-----BECK 2656  
 QY 399 VQOIEE-NSSEKRYKYVGNVETDINSSVDQKFPYAEISDLKREBELOQPIGAVDFLSQDD 457  
 Db 2657 VLBAESMHDVSVVKV-----STVNHLEFL-----RDELEERREKRRKQOABEE 2701  
 QY 458 SVEVNDSGVSESESEHQTPDDHLSPTASPYCRRYSDETPQSOOSGRTSGSDPGIC 517  
 Db 2702 TK-ITDKGKEDKKNRD-----SAGCTASNSNDTEQN-LSDGTPMDSYTTTP 2748  
 QY 518 SNTDSTQOAVLAKKRLKKAETELSDLYVSDKKOM-SPPFICEETDEQKLOTLDIGN 576  
 Db 2749 SSTDAATSE-----SKET--LGLTQSSQOQPLTPRPALGAEVQOELQSPAGEGSS 2797  
 QY 577 LEKAKLENSLSLGRDPSPIKKTSLSPSKGIYGSRLDLAKKHASLRQTESPDA 636  
 Db 2798 TOL-----LMPVEBELGPRPGEBAFTQWELSPATTISLSPERAEDS 2842  
 QY 637 DRTTLNADHSSKIVOHRLLSROEBLEKARVLLQARPDALKAGNKHTNTATPFCNR 696  
 Db 2843 DALT-----AVSSQLEGSFMDTSSLASCTLEBAVEDTS-AAQSSS----- 2881  
 QY 697 QLSDDQDEERRQLRBRARQLAABASGVYMSLPSYGENAAEKLEKRSYASG----- 749  
 Db 2882 -----QPRAG---SSTPGDAPPAVAEVOGRSDSGESAQPRE 2915  
 QY 750 DENNDIETDNEIEPGEFVVGGEDELTNLEND--DTPENGSLVNLK----- 795  
 Db 2916 DSSPPASSESSSTRDSAVNALSGADSRGILBEPILPSTSEBEDPLAGISLEPGVDPSFLAA 2975  
 QY 796 -----LKKLLEVOQVYANSPSS-----AAQAVTESSEQDMKSGTEDLARTERL- 838  
 Db 2976 LPDDIRREVLONGLRIPRTKTPSTNSSAPAVNGPCTEVSPBELAALPRAIQSEVLA 3035  
 QY 839 QKTERFR-----NEVFE--SKDSTVRKTQLOSPSOYT----- 869  
 Db 3036 QORAEQORRELAQNASDPFMDVTFIQLTLPDLRBSVLEDMEDSVLAWMPDIAABAQ 3095  
 QY 870 ENRPEKKRQSIQBDTKKNEEKAITTEQKSEBEVL--NKGFDQTSQYV---VGEI 923  
 Db 3096 LRREQARQQLMHEHLFGHSITSALSALIRSPAFTSRISGNQVGYTRLAVRGQGTFOH 3155  
 QY 924 AALENEQKQIDTBAALVERKRLRYLMD-----TGRNTEEBEEMMQEWFMLV 968  
 Db 3156 GCGSSNNRSGSNDTLLRLGRLLDHEALSCVLVLFVDEPGLNLSRLHVRVLRNICYHA 3215  
 QY 969 NKNALIRRNQOLSLKEKHDLERRRYELNRELAMLAIEDWQTEAQKR 1018  
 Db 3216 QTRHWVIRSL--LSIIQSSSESE-----LCIETPPLTTSSEK 3250

RESULT 12  
 US-10-511-937-2597  
 ; Sequence 2597, Application US/10511937  
 ; Publication No. US20060088836A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
 ; APPLICANT: Wohlgemuth, Jay  
 ; APPLICANT: Fry, Kirk  
 ; APPLICANT: Woodward, Robert  
 ; APPLICANT: Ly, Ngoc  
 ; APPLICANT: Prentice, James  
 ; APPLICANT: Morris, MacDonald  
 ; APPLICANT: Rosenberg, Steven  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
 ; FILE REFERENCE: 506612000104  
 ; CURRENT APPLICATION NUMBER: US/10/511,937  
 ; CURRENT FILING DATE: 2004-10-19  
 ; PRIOR APPLICATION NUMBER: PCT/US2003/012946  
 ; PRIOR FILING DATE: 2003-04-24  
 ; PRIOR APPLICATION NUMBER: US 10/131,831  
 ; PRIOR FILING DATE: 2002-04-24

PRIOR APPLICATION NUMBER: US 10/325,899  
 PRIOR FILING DATE: 2002-12-20  
 NUMBER OF SEQ ID NOS: 3117  
 SOFTWARE: PatentIn version 3.2  
 SEQ ID NO 2597  
 LENGTH: 851  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-10-937-2597

Query Match 3.0%; Score 166; DB 6; Length 851;  
 Best Local Similarity 17.8%; Pred. No. 0.00052;  
 Matches 148; Conservative 146; Mismatches 284; Indels 252; Gaps 34;

QY 403 EENSSEKSYKGVNVTEDNNSVDOEKFYAELSDLKREBELQPPISGANDPISQDSVAVN 462  
 DB 84 BDSKALSOXKJIKKKOTYKKEIDSNVLEKKETK-----AN 121  
 QY 463 DSGVSESESEHQT-PDHLSPSTASPYCR-RTKSDTEPOKQSSGRTSGSDPDCISNT 520  
 DB 122 HKG-----SENKTDLNSIGIKMKTSCKFKIDSNISKQDSKERTQNKKEKKNIVQHT 176  
 QY 521 DSTQAQVILGKKRLKATLELSDLVSDKKKMSPPICETDE-----OKLTLDI- 573  
 DB 177 TDSS-----LEKQRTLDSTSEIV-----KSPRIECSKTRREMOSVQLMTTRDSD 223  
 QY 574 -----GSNLEKEL-ENSRSL-ECRSDPSPITKTSLSFTSKLGYSDLDLAKKH 624  
 DB 224 GYENSTDEMCCKALBESSESVSISGDESENEITSVGA-----SDDDSEDE 276  
 QY 625 ASLQTESDPADRTTLNHAHSSKIYQHR--LRSQBELKERRARVLEQARRDALPAG 682  
 DB 277 EEDDEDEDEDEDEDDKSDSGPLANGKNIEISSDEDDTADLPFBESGFEHAMEL 336  
 QY 683 NKNNTATPFCNQLSDQDEERRRQI-----REBARQIJA-----EASGVKSEL 730  
 DB 337 DKD-----APRADEITRLAVCNMDWRLKADLALFNSPKFGVTFSVK 383  
 QY 731 PSYEMAAEKLEKERSKASGDENDNIEIDTNEIPEGPVVGSGDELTYNEN----- 780  
 DB 384 IYSEFGKERKEKQOVQVPELISIP-----EDAPF-----KQWTSKELRDYQFKLK 432  
 QY 781 -----DLDTPEONSKL-----VDLKL-----KKLLEVQPVAN 808  
 DB 433 YTYAVVDCDSPELTASKIYEDDCGLFESSCSFIDIRFPDITDPDEKDVASEYNLTAY 492  
 QY 809 SPSSAQAAYTESSEBODKSGTEIDLRTERLOKTERFENPVVFSKOSTVRKTOLQSFQY 868  
 DB 493 KPKFTSAAGTSTVETIWDETH---ERITMLNRKFKCEBLMD-----PGAY 539  
 QY 869 I-----ENRPEMKRQ-----SIQDPTK-----KNEBEKA-----ITETQKPSD 905  
 DB 540 LASSSEDEBEIEBELQCGDGVNVEDGKTKKQKQDEBQIAKYOQLQVIOKEKEKREN 599  
 QY 906 EVLNR-----GFKDTSQVVGELALENEKQIDTRALYERKRLRYLMDTRNTEERAM 960  
 DB 600 DMESEIKVVPOLKESASBEVKN--KLBSGDKLTPEOFLKKEKCKGRKRRQKALAEAS 657  
 QY 961 MQE-----WV-----MLVKNKALIRMNQSLLEKEHDLERREYELNRELPRAM 1004  
 DB 658 EEELPSVDLNDLPFAEBEVQIGINKS--VKSADGTSPEBEIEIRQ-----KAEWMLT 711  
 QY 1005 LAIED-----WQTEBAQRRQQLLDELVALVN----- 1032  
 DB 712 MMDEDEBSKGFNNYKIVEHQNTLSKKKKQKQMLKKKELLIEDFEVAVVNDARFOAMYSHLF 771  
 QY 1033 -----KRDALVRDLDAQKQABEDDEHLERTLEONKGRKAKKEK 1072  
 DB 772 NLDSDPNFKTKYAMEKILBEKAROREKEQELTOAIKKESEIEKESQ 821

RESULT 13  
 US-10-196-749-346

Sequence 346, Application US/10196749  
 Publication No. US20060094864A1  
 GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
 APPLICANT: Chen, Jlan  
 APPLICANT: Deanoyers, Luc  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Goddard, Paul J.  
 APPLICANT: Gunney, Austin L.  
 APPLICANT: Pan, James  
 APPLICANT: Smith, Victoria  
 APPLICANT: Watanabe, Colin K.  
 APPLICANT: Wood, William I.  
 APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ACIDS ENCODING THE SAME

FILE REFERENCE: P3430R1C340  
 CURRENT FILING DATE: 2002-07-16  
 PRIOR APPLICATION NUMBER: 10/052586

PRIOR FILING DATE: 2002-01-15  
 PRIOR APPLICATION NUMBER: 60/059263

PRIOR FILING DATE: 1997-09-18  
 PRIOR APPLICATION NUMBER: 60/059266

PRIOR FILING DATE: 1997-09-18  
 PRIOR APPLICATION NUMBER: 60/062250

PRIOR FILING DATE: 1997-10-17  
 PRIOR APPLICATION NUMBER: 60/063120

PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063121

PRIOR FILING DATE: 1997-10-24  
 PRIOR APPLICATION NUMBER: 60/063486

PRIOR FILING DATE: 1997-10-21  
 PRIOR APPLICATION NUMBER: 60/063540

PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063541

PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION NUMBER: 60/063544

PRIOR FILING DATE: 1997-10-28  
 PRIOR APPLICATION data removed - See File Wrapper or PAM.

NUMBER OF SEQ ID NOS: 612

SEQ ID NO 346

LENGTH: 671

TYPE: PRT

ORGANISM: Homo Sapien

US-10-196-749-346

Query Match 3.0%; Score 164.5; DB 6; Length 671;  
 Best Local Similarity 22.0%; Pred. No. 0.00047;  
 Matches 145; Conservative 102; Mismatches 260; Indels 151; Gaps 31;

QY 465 GVGSESEHQTDPDHLSPSTASPYCRRKTSDEPOKQSSGRTSGSDP--GIC----- 517  
 DB 76 GFNEGLWEIQN-NPHASYAVPPV---SSDSSEAPBANPADGDAEDDEDRGVAVATVAV 131  
 QY 518 -----SNTDSTQAQVILGKKRLKATLELSDLVSDKKKMSPPICETDEQ 567  
 DB 132 TATAASDRMSDSDDSDSNGLKR--KTPALMS--VSKARKAS-----SDLQAS 181  
 QY 568 LQTDIGSNLEKELTENSRLSECRDPE-SPIKKTSLSPTSK--LGYSDRLDLAKKH 624  
 DB 182 V-----SPSESESSSESEKTSDDPTPEKKAARARPRGLG-----GRKKKA 228  
 QY 625 ASLQTESDPDAD-----RTTLNHAHSSKIYQHRLLSRQBELKERARVLEQAR 674  
 DB 229 PSADSDSKADSDAKPEPVAMARASASSSSSSSDSVSVKPPRGRKPAEKPLPKPR 288  
 QY 675 RDAALTKGNNTTATPFCNQLSD-----QDBERRQRLREBARQILAEAS 723  
 DB 289 -----GRPKPE-RPPSSSSSDSDDEVDRISWTWRDARRRREARRRQER--- 337  
 QY 724 GVKMSELPSYGEAAEKLERSKASGDENDNIEIDTNEIPEGPVVGSGDELTYNLENDLD 783

Db 338 -----EIRLRLEDEKEE-KERRERADRG-----EAERSGSGSSGDEL-EDDER 380  
Qy 784 TPEONSKVLVDLKLKLELEVOP--QVANSPPSAQKAVTESSEQOMKSGTEDELFTERLOKT 841  
Db 381 VKRGRGRGRGPRSSSDSEBEALEREAKKSAKPPSSSTEPARKGQCKEVRPEKQ 440  
Qy 842 TERPRNFVFSKOSTVARTKLOQSFQYIENRPEMKRQSIQEDTYKKNER-KAALITETOR 900  
Db 441 Q-----AKPVKVRTRKRSEGFMDRKEVKKCKEVSVEKLOKHSEIKFALKVD-- 489  
Qy 901 KPSEDEVLNKGFMDTOSQYVVGELALENE-----OKQDTRALVKKRLTYLMDGRTTEB 956  
Db 490 SPVUKCLN-----ALBEETGLQVTSOILQKNTDVATL-KKLRRYKANDVNEKA 539  
Qy 957 EEAAMQEFMLVKNKALIRBMNQLSLEKEDLEREYELLNRELAMLAIEDMOQTEA- 1015  
Db 540 AEVYTRAKSRVLPKRTIAVQKVKAG-MEKKEKAER-----LAGEBELAGEAP 586  
Qy 1016 QKREOQLLDELVALVNRKDALVRDLDAQEKQAEDEHLERTLQONKGMKAKKEKC 1073  
Db 587 QKRAEDKPTDLSAPV-----GEATSOQGESAEDEKHEB-----GRDSEEGPRC 631

RESULT 14  
US-11-252-276-22  
; Sequence 22, Application US/11252276  
; Publication No. US20060094868A1  
; GENERAL INFORMATION:  
; APPLICANT: Giuliano, Kenneth A.  
; APPLICANT: Bright, Gary  
; APPLICANT: Olson, Keith  
; APPLICANT: Burroughs-Tencza, Sarah  
; TITLE OF INVENTION: A System for Cell Based Screening  
; FILE REFERENCE: 97-022-K2-CO  
; CURRENT APPLICATION NUMBER: US/11/252.276  
; PRIOR FILING DATE: 2005-10-17  
; PRIOR APPLICATION NUMBER: 09/713,572  
; PRIOR FILING DATE: 2000-11-15  
; PRIOR APPLICATION NUMBER: 09/430,656  
; PRIOR FILING DATE: 1999-10-29  
; PRIOR APPLICATION NUMBER: 09/398,965  
; PRIOR FILING DATE: 1999-09-17  
; PRIOR APPLICATION NUMBER: 09/031,271  
; PRIOR FILING DATE: 1998-02-27  
; PRIOR APPLICATION NUMBER: 08/810,983  
; PRIOR FILING DATE: 1997-02-27  
; PRIOR APPLICATION NUMBER: 60/136,078  
; PRIOR FILING DATE: 1999-05-26  
; PRIOR APPLICATION NUMBER: 60/106,308  
; PRIOR FILING DATE: 1998-10-30  
; NUMBER OF SEQ ID NOS: 168  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 1610  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:  
; OTHER INFORMATION: EYFP-DEVD-MAPA-BBFP construct  
US-11-252-276-22

Query Match 3.0%; Score 164; DB 7; Length 1610;  
Best Local Similarity 18.3%; Pred. No. 0.0018;  
Matches 186; Conservative 127; Mismatches 404; Indels 298; Gaps 37;

Qy 4 YASPMFTQTDVKKIKFPLSKKVSAALQFSLSCIFLRGKATBDMQSLASLVSMKQADI 63  
Db 204 YOSALSKDNEKRDHNVLLFVTAAGITLGMDELY-KKGDEVD-----GMADL 250  
Qy 64 GNLDDPEEDNEDDERVNOEKAQK-----ITELINKNFLDEAKQDLATVNSNP 115  
Db 251 SLVDALTEP-PPETEGIKRDFMAALBAEPYDIDVGEIVTEKTEFILLDDERTGNSSEK 309

Qy 116 DDP--DAEALNPGDDPDE-----EPITETASP-----RKTEDESPYNNSYNPFK 157  
Db 310 KKCCLDTQYVE--GIPSSKPTLLANGDHMEGNNTAGSPFDPLBBRDYDYOSSQWMP 367  
Qy 158 EVQ----TPQYLANPDEEAVYTKDSEPPQSTKKNIRPVDMSKLYADSSKTEBEELDES 214  
Db 368 DAEFCFQPOQVLDLDOAEFPNEHRDDGLADLLFVSSGPTVASAFTERDNPSEBSYGMPLC 427  
Qy 215 NPFYE-----PKSTPPNNLVNPFQELETERVRVKAAPAPVLSKPTGV----- 258  
Db 428 DSPASTAVUSQEWVAGAPNSPCSESCVSPETVETTLQAPATELSKAAEVSVKQOLPALKAL 487  
Qy 488 ETNABQTTDVVHSPSTDTTGPTBALAND-----IEIRK----- 524  
Db 317 WRNGLSPCALHHPRLDLIDYKSLNPODIKENNKAAYDGFASIGISRLRP---SDMWLL 373  
Qy 525 -----PDVI-----LANVTPSTESDMFLA 544  
Db 374 AIPDKLTWMTYLYQIRAHFSGEL-----NVVQIEENSSKSTYK-----GN----- 415  
Db 545 QDMELLT-----GTEAIAHANNITLPTBEDSSTKQVADPMESEIYVGNDDTS 591  
Qy 416 -YETDNNSSVDQEKFYAEISDLKREBELQOPISGAVDFLSQDDSVFVNDGVE----- 468  
Db 592 PKTEETTLPIKMDLAPBEDVLT/KETEL-APAGWVSLSIEEALAKNDVRSALPIVAQE 650  
Qy 469 ---SESEHQTPDDHLSPSTNAPYCRRTKSDTBEOQOQSSGRTSGDDPGICANTDSTQA 525  
Db 651 TVVSETEVVLATVAVLPS--DPYTLTLKQVTLPLBAER-----PLVDTMTSLST 698  
Qy 526 QVLAGKKRLKARTLELSDLYVSDKKKMSPPICET-----DEOQLQTLDIGSNL 577  
Db 699 EMTLAKETAPPTET-----NLGWA--KDMSELPSEVTLAKQVYLLPETKVAEPNVTPL 751  
Qy 578 EKEIKENSRSLBCRSDPESPRIK-----TSLSPYSKLGYSYSRDLDAKKH 624  
Db 752 SEEHVTSVDMQSPSAETBAVPLAKNAULHSGTELVIVNSMAPADDLPLETKVATVPYKD 811  
Qy 625 ASLRQTESPPDADRTTLNHAHSSKIVQHRLLSROBELKERARVLLQARRDALQAGNK 664  
Db 812 KGTVOETEBR-----REDSQL-ASMQ 831  
Qy 685 HNTNTATPCNRQLSDQOQBERRQLERABQLIAEARGGVKMSLPSYGEMAABKLKER 744  
Db 832 HKGQSTVPCTASPP-----PVKABQKSTLPIDAPSPLENLBOK 871  
Qy 745 SKASGDENDNIB-IDTNEBIPBGVVGDDDELTNLENDDTP--EQNSKLVLDLKLKLE 801  
Db 872 ETPGSPSEPCSGSVSRQGEAKAAVGTG-----NDITTPNKKEPPSPBKKAPLAT 923  
Qy 802 VQPVANSPSSAQAQKAVTSSSEQDMKSGTEDELFTERLOKTTTERFNPPVVFQSKDSTVRKTO 861  
Db 924 TOP--AKTSTSKAKTQPTSLPKQAPPTTGGAL-----KKPMSLASGVPAAPI 970  
Qy 862 LQFSQYIENRPEMKRQSGIOEDTKKGNBEKAALITETQ-RKQSEDEVLNKGFMDT 915  
Db 971 KRPAAATATAPRSTLPARDVK--PKPITAKVAEKRTSPSKPSASPALKPGPKTT 1023

RESULT 15  
US-10-511-937-2566  
; Sequence 2566, Application US/10511937  
; Publication No. US20060088836A1  
; GENERAL INFORMATION:  
; APPLICANT: EXPRESSION DIAGNOSTICS, INC.  
; APPLICANT: Wohlgenuth, Jay  
; APPLICANT: Fry, Kirk  
; APPLICANT: Woodward, Robert  
; APPLICANT: Ly, Ngoc  
; APPLICANT: Prentice, James  
; APPLICANT: Morris, MacDonald

```

; APPLICANT: Rosenberg, Steven
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING
; TITLE OF INVENTION: AND MONITORING TRANSPLANT REJECTION
; FILE REFERENCE: 506612000104
; CURRENT APPLICATION NUMBER: US/10/511,937
; PRIOR FILING DATE: 2004-10-19
; PRIOR APPLICATION NUMBER: PCT/US2003/012946
; PRIOR FILING DATE: 2003-04-24
; PRIOR APPLICATION NUMBER: US 10/131,831
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: US 10/325,899
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 3117
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2566
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-511-937-2566

```

Query Match 3.0%; Score 162.5; DB 6; Length 1186;

Best Local Similarity 18.4%; Pred. No. 0.0015; Matches 221; Conservative 167; Mismatches 417; Indels 395; Gaps 53;

```

QY 2 KOVASPMPTOTDVKLKFKPKSKVVSAALQPSLCIFREGKATDEDMQSLASIVSMKQA 61
DB 95 KOLASSDSRKTTEKLTFTLKQAIKTAFR-----SKRDEALPSLTQV--RREN 141
QY 62 DIGNLDPEED-----NEDDDE--NRVQEBKAKITELINKNLFDEAKDLATVNSN 113
DB 142 DLYLPLQEBEKSSBEDEKEMQERNQ--KQALQEFHNPQALIDISEDPSSL---- 196
QY 114 PFDDPDAAELNPFQDPSEBPTF-----TASPKRTED----- 146
DB 197 ---PPEV-----KHRIITDKEFTKRRRTLFKAMPESBDPFGYQKGLKKNYL 243
QY 147 -----SFYNNSYNPFKEVOTPOYLNPFDEBEAFVTKDSPPOSTK 186
DB 244 NQHIENHVKEMNQOHSIHRIQYEDGFLKESRRVVS--EDTSYILIKGIOAKTVA 301
QY 187 RKNIRPVMSKYLYADSSKTEEBELDESPPYEPKSTPP--PNNLV-----NPVQEL 236
DB 302 EVDSESLPSSSKMGMSFDVSSPCBKLTKEBPDPATPPRILLAMQALLGSSSEEL 361
QY 237 ETERKRVK--RKAPV-----PPVLSPTKGVLEN--TVSAGKDLSTPKSPPIPS 282
DB 362 ESENRRQARGHNAAPAAVDEGISPRTLGAIKRALDDEDDVACAGDDVQGT----- 412
QY 283 VLGKRPNASQSLVWCKEVTKNY--RGVKITNFTTSMRNGLSFCALIHFRPDLIDYKSLN 341
DB 413 ---GPGAEKRT--NSSTENSDEGLKVRD---GKGIPTATLAS-----SSVN 452
QY 342 PQDIKENNKAYDGFASIGISRLLEPSDMVLAIIPDKLTWTVYLYQIRAHFSGOELNVQ 401
DB 453 -----SAEHHVASTNEGR--EPTDSV-----PKGMSLVH 480
QY 402 IENSSKSTYVGVNVTDTNASSVDOEKYAEISDKREPELQOPIGAVDPLSODDSYFV 461
DB 481 -----VGTFAF--PISD-----ESMKDKRDLPLESAVVR 509
QY 462 NDSGVGSESEHQPDDHLSPTASPYCRKTKDTEPOKSOSSGRTSGSDDPGICSNTD 521
DB 510 HSDAPGLNGRELTTP-----ASPTCTNSVSKNE----- 537
QY 522 STQAVLLGKRLKKAETLALSDLYVSDKKOMSPFICEETDQKQLDLDIGSNLKEK 581
DB 538 -THAEVLEQNELCYBESKPDSSILSSDDTKCKENSASVYIGPVSLOETSIYSVPSBA 596
QY 582 LENSRSLECRSDPSPPIKTSLSPTSGLGYSYRDLIAKKHNASLQTESDPPADRTTL 641
DB 597 VDNVENV-----VSFNAKEHENFLETTIOEQTTESAG 628
QY 642 NHAHSSKIVQHRLLSROBELKERARVLLBOARRDAALKAQNKNTNTATPFCNRQLSDQ 701

```

```

DB 629 QDLISIPKAVPMEIDSESDSPFIEVQSVISDEELQA-----EPFETSKPSEQ 680
QY 702 QDEERRKQLRERAKQLAENASGYMSELPSYGEAAEKLEKRSKASDENDNIFIDPNE 761
DB 681 GEER-----LV-----GTREGAPAESSSLARDNSBRDVEDG--EPOEAKQAD 723
QY 762 EIPGPFVVGGDDELTNENDLDTPEONSKVLDLKLK-----LLEVO----- 803
DB 724 SLHEWQDI--NLEELTEESNL--LAQONS---LQAQKQOERIAATVYGQMFLESQELLR 777
QY 804 ---POVANSPPSAQAQAVTESSEQDKMSGTEDLRTERLQKTERFRNPVVSQDSTYRK 859
DB 778 LFGIPYIQAPWEAEQAQALIDLTDQTSGETTD--SDIWLFGARHYVRN--PFNKXKPF-- 832
QY 860 TQLOSPQYIE--NRPEKQKRSIQ-----EDTKKNEERKALITETQKSEBDEVLANK- 910
DB 833 ---EYQYVDYFNQGLIDRNKCLINLAVYLGSDYTEGIPYVGCVTAM-----ELNBF 881
QY 911 ---GFKDTSOVVVGELALLENBOKQIDTRALVYERKRLRYLMDTG--RRTSEBAMQOEW 965
DB 882 PCHGLEPLKFSWMHBAQKPKIRPNPDQTKVKQLTQLTGFFNPVAVAEALYK-- 939
QY 966 MLVNNKKNALIRRNQQLSLEKEHDLERRRYELNRELPAMLAIEDWQKTEAQKREBQLILD 1025
DB 940 -VVDSDSKG-----SFLWGPDLDDKIREFCQRYF-----GWNRT--KTDESLEF-- 978
QY 1026 ELVALVNRDALVDDLDAQEKQ-----AEEDEHLEFTLBONKQK---NAKKEK 1072
DB 979 -----FVLKQLDAQOTQLRIDSPFRLAQOKEKEDAKRIKSQRLNRAVTCMLRKEKB 1028

```

Search completed: May 20, 2006, 17:30:35  
Job time : 18 secs

**THIS PAGE BLANK (USPTO)**







Qy 1051 EDEHLERTLEQNKGMKMKKEK 1072  
Db 1014 LOSQLENRQOQSQNSQLEKDDQ 1035

RESULT 3

528381  
utrophin - human  
N:Alternate names: dystrophin-related protein  
C:Species: Homo sapiens (man)  
C>Date: 17-Apr-1993 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004  
C/Accession: S28381, S28914, S03966  
R:Tinsley, J.M.  
submitted to the EMBL Data Library, November 1992  
A/Reference number: S28381  
A/Accession: S28381  
A/Molecule type: mRNA  
A/Residues: 1-3433 <TIN1>  
A/Cross-references: UNIPROT:P46939, UNIPARC:UPI0000173E6A, EMBL:X69086, NID:G34811, PIDN  
R:Tinsley, J.M.; Blake, D.J.; Roche, A.; Fairbrother, U.; Rise, J.; Bych, B.C.; Knight,  
Nature 360, 591-593, 1992  
A/Title: Primary structure of dystrophin-related protein.  
A/Reference number: S28914, MUID:9306045, PMID:1461283  
A/Accession: S28914  
A/Molecule type: mRNA  
A/Residues: 27-246;2839-3343 <TIN2>  
A/Cross-references: UNIPARC:UPI0000173E73, UNIPARC:UPI0000173E74, EMBL:X69086  
R:Love, D.R.; Hill, D.F.; Dickson, G.; Spurr, N.K.; Bych, B.C.; Marsden, R.F.; Walsh, F.  
Nature 339, 55-58, 1989  
A/Title: An autosomal transcript in skeletal muscle with homology to dystrophin.  
A/Reference number: S03966, MUID:89238543, PMID:2541343  
A/Accession: S03966  
A/Molecule type: mRNA  
A/Residues: 2944-3433 <LOV>  
A/Cross-references: UNIPARC:UPI000000673, EMBL:X15488, NID:G30933, PIDN:CAA3515.1, PID  
C/Comment: This protein is found primarily at the neuromuscular junctions in adult muscle  
d regenerating muscle.  
C/Genetics:  
A/Gene: GDB:UTRN; DMDL  
A/Cross-references: GDB:119851, OMIM:128240  
A/Map position: 6q24-6q24  
C/Superfamily: dystrophin; alpha-actinin actin-binding domain homology; spectrin/dystrog  
C/Keywords: actin binding; cytoskeleton; leucine zipper; membrane-associated protein; mu  
F:30-248/Domain: alpha-actinin actin-binding domain homology <ACT>  
F:308-417/Domain: spectrin/dystrophin repeat homology <SP01>  
F:418-526/Domain: spectrin/dystrophin repeat homology <SP02>  
F:528-637/Domain: spectrin/dystrophin repeat homology <SP03>  
F:638-685/Region: hinge  
F:686-796/Domain: spectrin/dystrophin repeat homology <SP04>  
F:804-902/Domain: spectrin/dystrophin repeat homology <SP05>  
F:906-1013/Domain: spectrin/dystrophin repeat homology <SP06>  
F:1015-1122/Domain: spectrin/dystrophin repeat homology <SP07>  
F:1124-1230/Domain: spectrin/dystrophin repeat homology <SP08>  
F:1232-1334/Domain: spectrin/dystrophin repeat homology <SP09>  
F:1339-1450/Domain: spectrin/dystrophin repeat homology <SP10>  
F:1451-1541/Domain: spectrin/dystrophin repeat homology #status atypical <SP11>  
F:1543-1649/Domain: spectrin/dystrophin repeat homology <SP12>  
F:1651-1755/Domain: spectrin/dystrophin repeat homology <SP13>  
F:1856-1973/Domain: spectrin/dystrophin repeat homology <SP14>  
F:1975-2081/Domain: spectrin/dystrophin repeat homology <SP15>  
F:2083-2185/Domain: spectrin/dystrophin repeat homology <SP16>  
F:2227-2333/Domain: spectrin/dystrophin repeat homology <SP17>  
F:2335-2440/Domain: spectrin/dystrophin repeat homology <SP18>  
F:2442-2556/Domain: spectrin/dystrophin repeat homology <SP19>  
F:2558-2688/Domain: spectrin/dystrophin repeat homology <SP20>  
F:2690-2797/Domain: spectrin/dystrophin repeat homology <SP21>  
F:2798-2869/Region: hinge  
F:2812-2849/Domain: ww repeat homology <RW1>  
F:2837-3117/Region: cysteine-rich  
F:3263-3284/Region: leucine zipper motif  
F:3328-3349/Region: leucine zipper motif

Query Match 5.9%, Score 326.5; DB 1; Length 3433;

Best Local Similarity 19.8%; Pred. No. 6.4e-06;  
Matches 197; Conservative 159; Mismatches 336; Indels 303; Gaps 37;

Qy 286 RKPAASGLLWCKEYVTKNRYGVKTNFTTSMRNGSFCALIHFRPDLIDKSLNPQDI 345  
Db 148 QQTNSKILSWRQOTRYPYSQVAVNLFTTSDGLFANLVLRHKEDFLSWDKVYKMS 207  
Qy 346 KENKKAADGFAS-IGISRLLEPSDWVLAIPDLITVITYLOI- 388  
Db 208 IERLEHAFSKAQYTLGIEKLLDPED-VAVALPDKSLITWLTSLFEVLPOQVITDAIREV 266  
Qy 389 ----RAHFSQELNVQVIEENSSKTYKVGNFTD---TNSVDQEFYALSDLKBBP 441  
Db 267 ETLPRKYKCEBEALINIQSTAPEEHSBRATPTSTVTEVMDLDSQYALBEV----- 321  
Qy 442 LQPIGAVDFLGGDSVFVNDGVSSEHOPPDHLSPTASP-----Y 488  
Db 322 LTWLISADTFQEDD---ISDD-VEEVKQFATHEAFMEELTAHQSSVGSVLQAGNOLI 377  
Qy 489 CRATKSPTEPQKSG-----QSGRTSGSDPGICSNSTDSTQAVL----- 528  
Db 378 TQGTLSDEEFELQEQWTLNARWEALRVSMRQSLHD--VLMELQKKQLQQLSMILT 435  
Qy 529 LGKKRLKAETLLESD-----LYVSDKKQKMS 556  
Db 436 LTBERRIQMETCPDLDVKSLOKLEBKSLQSDLEABQVNSLTTHWVIYVDENSGESA 495  
Qy 557 PFI-----CETDEQ-KLQTLIDGSLLEKLENSLSLCSRSPSEPIK 599  
Db 496 TALIEDQLQKGRWTAVCWTEBRMVRLOEINI--LMQELLEE--CCLKAMLTLEK 548  
Qy 600 KTSLSPTSKLGYSGYSRDLDAKKKHAFLROTSPDADRTLLNADHSSKIVQ----- 652  
Db 549 BEALNKQTSNFKDQKSLSVRLALKE--DMEKRQTL--DLSLIGQVQGLD 602  
Qy 653 ----HRLSROBELKER-----ARVLQEAR-EDAA 678  
Db 603 NSKASKKINDSEELTQRMWSLVQRLBDSNQVTAQAVKLGMSQIPQKDLLETVRVAEQ 662  
Qy 679 LKAGNKNTTATPFCNRQLSDQDERRRQLERRAQL--IAEASGYMSLSLPSYGE 735  
Db 663 ITTKSKQELPPPPPKRQI--HYDIEAKKFPALISELNLWKMTAQTIEIKYMK 720  
Qy 736 MA-AEKIKERSKASGENDNIBIDYNEIEPGFVVGSGDELTLNENDLDPQNSKLVLD 794  
Db 721 MQDTSEKKKKA-----LEKQERLIR-----ADBLNQTGQLVGEQKSGILPTE 767  
Qy 795 KLKLELVQPVANSPSSAAQKAVTESSEQDMKSGTEDLRTERLOKTERFRNPFVPSKD 854  
Db 768 EIKNVLE-----KVSEEMKNVSGHLEDLRK-----IQAQED 799  
Qy 855 STYKQTQLOGSPSOYIENRPMKQKRSIQEDTK-----KGNBEK 892  
Db 800 INNVFKQDLDEKVIKTKKEWVGHGTSISRSQSLPSLKSCQRELTNLGLHPKIMAR 859  
Qy 893 AATTEQKRSSEBVLNKG-----KQTSQYVGEALLENQKQIDIRAA 938  
Db 860 ASGALMSQPSADPFVQRGDFLGRYQVQAEVDEQDQ-----LENELKQPGHAY 912  
Qy 939 LVEKRLAYLMDTGRNT-----EEBAMQGEFMLVVKNAALIRRM 978  
Db 913 L--ETLKLTDVANDSNKQVSLANTLDAKYKALQEKKTIDE--ILNQGPAHLKLA 968  
Qy 979 NQSLLEK--EHDLERRYELLNBLRAMLAIEDWQKTBQAKRRBQLLDLDELV----- 1028  
Db 969 EETKALEKKNVHPVEKLYKQEPDVGQ-----KWNKLLKVLVSKDILLBEALTLTARAFA 1023  
Qy 1029 --ALVNRKDLVYDLDQAQKQAEEDHEHLERTLEQ 1061  
Db 1024 DSTVIEKMDGVDFLWKQQAQGDAGLRQLDDQ 1058

RESULT 4

T34418  
 hypothetical protein F12F3.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34418  
 R:Fulton, B.; Wohlmann, P.  
 submitted to the EMBL Data Library, July 1998  
 A:Description: The sequence of C. elegans coemid F12F3.  
 A:Reference number: Z21521  
 A:Accession: T34418  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-3488 <FUL>  
 A:Cross-references: UNIPARC:UPI000017B8E6; EMBL:U00022; PUDN:AC25885.1; GSPDB:GN00023;  
 A:Experimental source: strain Bristol N2; clone F12F3  
 C:Genetics:  
 A:Gene: CESP.F12F3.3  
 A:Map position: 5  
 A:Insertions: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 5.9%; Score 324.5; DB 2; Length 3488;  
 Best Local Similarity 19.1%; Pred. No. 8e-06;  
 Matches 245; Conservative 215; Mismatches 429; Indels 397; Gaps 54;

25 VVSAAQFSLSCTFLREGKATDEDMOSIASLVSMKQADIGNLDP---FEEDNED-DDENR 80  
 315 VAGEAWCFSDVVVHMGSE-ESRDD-----KSVDEVDSITVLEKKDGDGDKSK 361  
 81 VNQEEKAATTEILINKLFLDEAKCOLATVNSNPPDDPAEELNPGDPDSEPTETAS 140  
 362 PKKKKKTIKKKE-TPESEQVTALEPEQOKISEVDVQSVATEVGAKKKPAERF-TDL 419  
 141 PRKTEDSFVNSNVPKFEVQTPOLYNPDEBEAFVITKD-----SPPOSTKKNR 191  
 420 AKKDSK-----KKSDEPEASTEESTKTTKPTNDKTSKSAEKKTVK 461  
 192 PVDMSKYLVADESKTEEBELDSNPFYKSTPPNNLVNPVQELTEERRVKKAP---- 247  
 462 PKKEVVGKPLEAKCPEDKDKASQPSKSSKSSPP-----TDGKKKKQIKALE 509  
 248 APPVLSPTKVLNENTVSAKDLSTSPKSPSPVLAQRKPA-SQSLLVCKEVTKNR 306  
 510 IPDELSSRRG-----DPTVMSEFTNITTTIRGEGSADAKTPLY--EPLSASV 555  
 307 GVKITNFTTSMRNGSFCALIHFRPD-----LIDVKSINPQD-----IKE 347  
 556 SMKVFPLVESAKKAEFSFRKSETPDDSKRKEGPRPAKSEKQEVTAEKOSTALE 615  
 348 NNKKAVDGFASIGISRLLE--PSD--MYLLAIPDKLVTWTVLYQIRAFSGQEL--NVV 400  
 616 SKKKEVDE-----SKISQQPSDKKKSSEVGVPEKA-----AGETKKQDV 656  
 401 QIEBNSKSTYVGNVETDTNSSVDQEKYAEISDLKREBELQAPISGAVDFLSQDSVF 460  
 657 EIEEVPKKTITIKKTEKSD--SSISQKSNVLKPAD-----DDKSKSDV- 698  
 461 VNDSGGSESEHQEPDDHLSPSTASPYCRKTSQDTEPKSQSSGRTSGSDPGICNT 520  
 699 -----TDKSKTKTEDQTKVATDS-----KLEPAADTTKQLETEVVDK- 737  
 521 DSTQAVQLGKRLKAETLEISDLVSDKKDMSPFI-----CEETDEQKQTL----- 571  
 738 -----SKKKVLKKKT-EKSDSFISQSE--TPPVVEPTKPAASEKQAEVVKAKK 785  
 572 --DIGSNLEKEKLENSRSL--ECRSDPSPIKTSLSPTSCLKGYSDLDLAKKHAAS 626  
 786 QKEVDNLTREAEVAAKIADEKLKIEAANIKTA-----EVEAAKKQKQEK 832  
 627 LAQTESDDP-----ADRTTLNHADHSKIVQHLLSNOEELKEKARVLEQARDAA- 678  
 833 DEQLKLETVVSKSAAEKLEKQAIKKAELADVVKQKEINENK--LEAAKKSAD 890  
 679 -----LKAQNKHNTNTATPFCNRQL-----SDQDDE 704

Db 891 KKLSESAKSKKVESEVYKPEGEKTKRAGEKTVQVSESPFSKTIIDTKDVGATEPADE 950  
 Qy 705 ERRFQL---REPARQLAE--ARSGVQSE-----LPSYGMAAEKTKERSQASGD 750  
 Db 951 TPKKKTIKKTKESDISQKATDESEKSKQKODEPTKPAVSEFQMTVEADSKSKQKE 1010  
 Qy 751 ENDNITID-----TNEBIPGCVVGGSELTLENDLTPBEQNSLVD-LKL----- 796  
 Db 1011 TDEKLDIAIAKTKQAEDESKLDAQEKIKVSESD-DAAKKEKLENDLKLESEIATK 1069  
 Qy 797 -----KCLEVQPVANSPPSAQAQAVTSSSQ-----DMKSGTEDLRTE----- 836  
 Db 1070 KASADLKLEBQQAQKAAVEAAKKQKEMDEQLKDTEAASKKAABELEKQAIKK 1129  
 Qy 837 -----RLQKTERFRANPVFSKDSVTAKTQLQSQYIENRPEMKRQRSIQEDPKGN 889  
 Db 1130 AAGADAIVKQKQKLEDE-KNLTLEANKKSAAKTLK-----IEESAAKSKQVVEBQAKLDA 1181  
 Qy 890 BEKAATETQRPKSEDE-----VLNKGPK-----DTSQYV 920  
 Db 1182 QYAKTAKEQTKLEKDEKTSKESKETYDEKPKKVLKKTKEKSDSSISQSEISKTVV 1241  
 Qy 921 GEIALALENE-----OKQIDTPRAL-----VEKRLRYL--MDTG 951  
 Db 1242 ESAGSESESTQVADAARQKQKTEDEKQKLEBITAKGSADESKLEAEKKAABEVA 1301  
 Qy 952 RUTTEERAMQEWFMVLNKKAL-----IRRMQSLLEKHDLEFRYELIN--- 998  
 Db 1302 KQKKEDEQLKLDTEAASKKAABELEKQSHIKKAAEVNDVVKQKELEEKQLESEA 1361  
 Qy 999 -----BELRMLAIEDQKTEAQKRRQELLDELVALVNRDLVLRDLQ 1044  
 Db 1362 TKKADAEKTLBQKKAABELEIETQKEQKLAQSQSLBEDA---KQSAEKQKLESE 1417  
 Qy 1045 EKQAEDEDELERTLEONKGMVAKKE 1070  
 Db 1418 TSKQTEAPKESVDEKPKKGYLVKK 1443

RESULT 5  
 FAD0AA  
 alpha-actinin - slime mold (Dictyostelium discoideum)  
 C:Species: Dictyostelium discoideum  
 C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 09-Jul-2004  
 C:Accession: S00103; A29006  
 R:Noegel, A.; Witke, W.; Schleicher, M.  
 FEBS Lett. 221, 391-396, 1987  
 A:Title: Calcium-sensitive non-muscle alpha-actinin contains EF-hand structures and high  
 A:Reference number: S00103; MUID:87304850; PMID:3622778  
 A:Accession: S00103  
 A:Molecule type: mRNA  
 A:Residues: 1-862 <NOE>  
 A:Cross-references: UNIPROT:P05095; UNIPARC:UPI0000125099; EMBL:Y00689; NID:G7177; PIDN:C  
 J. Cell Biol. 103, 969-975, 1986  
 R:Witke, W.; Schleicher, M.; Lotsepelch, F.; Noegel, A.  
 A:Title: Studies on the transcription, translation, and structure of alpha-actinin in Dic  
 A:Reference number: A29006; MUID:86304574; PMID:3745276  
 A:Accession: A29006  
 A:Molecule type: DNA  
 A:Residues: 92-359, 'P', 361-500, 'T', 502-505 <WIT>  
 A:Cross-references: UNIPARC:UPI000016898B; EMBL:X04324; NID:G7202; PIDN:CAA27855.1; PID:C  
 C:Suprafamily: alpha-actinin; alpha-actinin; actin-binding domain homology; calmodulin rel  
 C:Keywords: actin binding; calcium binding; duplication; EF hand; homodimer  
 F:1-21-336/Domain: alpha-actinin actin-binding domain homology <ACT>  
 F:266-377/Domain: spectrin/dystrophin repeat homology <SP1>  
 F:386-493/Domain: spectrin/dystrophin repeat homology <SP2>  
 F:502-607/Domain: spectrin/dystrophin repeat homology <SP3>  
 F:616-717/Domain: spectrin/dystrophin repeat homology <SP4>  
 F:730-762/Domain: calmodulin repeat homology <EF1>  
 F:766-798/Domain: calmodulin repeat homology <EF2>

Query Match 5.8%; Score 321; DB 1; Length 862;

Best Local Similarity 23.3%; Pred. No. 1.9e-06;  
Matches 187; Conservative 98; Mismatches 262; Indels 256; Gaps 35;

```
QY NASOSLWCKEVTNKGKVTNTTNTSMRNGLSFCALIHHPRLIDYKSLNPDIEN 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 138 SAKELILMCRKTEGYRVRKGNHTSPQDLACALIHKRPDLINPDSLNDKXGN 197
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 349 NKKAVD-GFASIGISRLLEPSDMVLLAIPDKLTVMYTYQIRAHPSGQELNVQIEKSS 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 198 LQALFDIAEKELIDIPKMLDVEDMDLVAPDERSWTTYAQYTHHSAS-----RKAETNG 252
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 408 KSTYKVGNYETDYNSSVDQEKPYAELSD-LKREPELOQPIGAVDFLSGDDSVFVNDGCV 466
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 253 KQVGVN----LDTFMLRQTK-----SDYLKRAVELVQMIN-----DKQASLESDF 295
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 467 GES-FSEHQTPDDHLSPTASPYCRKTSPTPEQSGSSGRTSGSDPGICSNITDIOA 525
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 296 GDSIESVGSFNNAH-----KEYKTEKPEPKGQEVN-----ELEA 329
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 526 --QVLLGKGRLLKAAETLELSDLYVSDKKGMSPPIC-----BERTDQKQLTLDIGSL 577
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 330 IYNELQTKLRLIKR-----PFVAPAGLTPNEID-----STWSAL 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 578 EKEKLENSRLSCRSDESPPI-----KKTSLSPTKLGYSVSRDL-DLAKKXHA 625
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 EKAQGEHAELRIELKQKQIAVLLQKYNRLIKLENNATTKSVYLSGNETGDSITAYQA 424
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 626 SLRQTE-----SDPD-----ADRTTANH-----ADHS 648
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 425 KLKMLAEFDEGCQSLSEGSNDLLIILQITELNNGVPELTERKQTFPAQQWGVKSSA 484
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 649 KIYVHRLISROBELKERARVILLEQARRDAL-----KAGNKH-----NNTATPFCNRQ-- 697
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 485 ETYNTLAELERLQKLEDSLVEFPAKRAQLNWMIEADHDVFPDFINVSQVQVQELQEK 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 698 ----LSPOOD-----BERRQ 709
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 545 FDAFLHDQSQFALEALALATQQLRELGRSENDYVSIYDELAKMNNLLAGIEERKQV 604
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 710 LREARQILAEARS-----GYKMSLPSTGYEMAELKSKASGENDN----IETD 758
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 605 L--ANELLTYTNNVLCQSPSVKANEISDYRVYTLDAISQTSDDPOEQLNTRAITTA 661
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 759 TNEBIEGFVVGGGDELNLNENDL-----DTPENGSLVDLKLK-----KLEVGQPOVA 807
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 662 HAERKPE-----LDELITTSQLEBAQVNDKHTQISLSLKLKMDKLTAKKNEGV 715
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 808 NSPSSAAQ-KAVTESSEQDMKS-----GTEDLRTERLQ-----KTERFRNPV 850
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 716 BGEILAQULGVTALELSEFACFSHPKDNNDKLNRLREFSSCLKSIGDELTEQLNQVI 775
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 851 PSKDSTVRKTQLOQSPQYIENRPEMKQRSIQEDTKG--NBEKAITETQRPSEDE 906
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 776 SKITDNGT--ISFEERIDYVSSRKGTDSVESTKAFAKVMADKDPITBAQ----- 826
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 907 VLNKGFQDTGQ--YVGEELALE 927
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 827 -IRRAISDSKQIDYLLASMPAVE 848
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 6

beta-spectrin - mouse  
C/Spect: Mus sp. (mouse)  
C/Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 13-Aug-1999  
C/Accession: 152577  
R/Bloom, M.J.; Birkenmeier, C.S.; Barker, J.E.  
Blood 82, 2906-2914, 1993  
A>Title: Complete nucleotide sequence of the murine erythroid beta-spectrin cDNA and the  
A/Reference number: 152577; MUID:9403578; PMID:8219233  
A/Accession: 152577  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: mRNA

A/Residues: 1-2128 <RES>  
A/Cross-references: UNIPARC:UP10000029884; GB:S66283; NID:g440899; PIDN:AA828600.1; PID:  
C/Genetics:  
A/Gene: Spnd-1  
A/Introns: 418/3; 1742/2  
C/Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectri  
C/Keywords: actin binding  
F:53-271/Domain: alpha-actinin actin-binding domain homology <ACT>  
F:1787-1892/Domain: spectrin/dystrophin repeat homology <8PH>

Query Match 5.8%; Score 319; DB 2; Length 2128;  
Best Local Similarity 19.7%; Pred. No. 7.3e-06;  
Matches 221; Conservative 160; Mismatches 382; Indels 358; Gaps 44;

```
QY 285 GRKE-NASOSLWCKEVTNKGKVTNTTNTSMRNGLSFCALIHHPRLIDYKSLNPQ 343
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 169 GREQSAKDALLMCCQKTAGYPHVNATNFTSSWKDGLAFNALIHKRPDLIDPKLDS 228
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 344 DIKENKKAYD-GFASIGISRLLEPSDMVLLAIPDKLTVMYTYQIRAHPSGQELNVQ- 401
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 229 NARHNLHAFVVARQGIIFLDPED-VFTENDEKSIITTYVAHFHYFSKMVLAVER 287
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 402 -----IEE-----NSSK-----STY 411
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 288 KRVGKVIDHALETEKMIKYSGLASDLLTWIQTISVLSRKFANSLSGVQQLQAFSTY 347
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 412 KV-----GNYET-----DTNSSVDQEKPYAELSD-LDLKR-----EPEDQOP 445
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 348 RTEKPPKFOEKGMLEVLLFTIQSRMRANQKVTPPHDGKLVSDINRMESLEAEAYQRE 407
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 446 ISGAVDFLSODSVFVNDGCVG--SESEHQTPDDHLSPTASPYCRK-----SDTE 497
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 408 LALSELIRDE--FDKKAARETWLNENQVLTVQDNFGYDLAVEAKKHAEIETITA 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 498 PQKQ-----QSSRTSGSDPGICSNYDS-----TQAVLLGKGRLLKAELELSD 544
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 465 AYERVRVALLDLAQLERENYHDKRIIARDNILLMSYQLRLSRQLEBATLALQK 524
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 545 LYVS-----DKKMSPPICEPTDEQKLTQTLDGSLNEKLEKLENSSLERSPSESTY 599
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 LFQMLHSIMWDEIKHILISAEFGKILLEVEDI--LQKRLMEA-DIALQSGKVAAIT 580
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 600 KTSLSPTSKLGYN-----YSRDLDAKKKAGSLRQTES----- 632
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 581 AATIQFABGKGYQCDQVYIDRVSHLEQCPSELNNAAKRAKQLEBSKLMKFWENDE 640
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 633 -----DPDADRTLLNHADHSKIVQHRLISROBELK----- 664
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 641 AEWIKERKQIYSLDYGKDLTSVLIQKHKAFEDBLRGIDAHLKQIFQERADDMVAQKQ 700
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 665 -----PARVILLEGARD--AALKAGNKHNTTATPP-----CNRLSD 700
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 701 FGHPIQETRVKVSQAQMDHKEILAFAKKQLQDLENFFQFGADDDLKAWLQDAHRLISG 760
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 701 Q---ODEBRRQLAERARQLAERAGVNRKSELPSYEMAELKERSKASGENDNIEI 757
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 761 EDVGQDEGATRALGKKHKEFLLEESRGVNE--HLEHQAGPPEERFRSDVTNRLOA 817
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 758 -----DTNEBIEGFVVGSGD-----ELTNLENDLDT 784
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 818 LRKLYQGVLLQALRGHKLQLEALDLYTPGESDCELMATEKGMLOMDIPNLEBDEV 877
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 785 PEONSKLVLDLKLKLELVQPOVANSPESSAAQKAVTESSEBDMKSGTEDLRTERLQKTER 844
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 878 VQHRFDILDOBMKTLMAQIDGV-----NLANNLVESGHP--RSGEYQVQYDRLNKXWQA 930
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 845 FRNPVYVSK---DSTVKK-----TQLOQSPQYIENRPEMKQRSIQEDTKKGNBEKAITB 897
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 931 FOAVVSGRLEAVDALKVNNYTCVDCSTSKMIDKTK-----VVESTKQLGQDLAIGYA 984
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 898 TORKPS--EDBEVLN-----KGFQDTSQYVVGELALLENBQKQ-IDTPAALVEKRLRYLMD 949
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 985 IQRLSLGLERLDVLAIRVRGALBRESGY-----LMESHPEQKEDICGQADVYKLMKGLQD 1040
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```



R. Byers, T.J.; Brandin, E.; Lue, R.A.; Winograd, E.; Branton, D.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 6187-6191, 1992  
 A>Title: The complete sequence of *Drosophila* beta-spectrin reveals supra-motifs comprised  
 A:Reference number: A46147; MUID:92335263; PMID:1631106  
 A:Accession: A46147  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-2291 <BY2>  
 A:Cross-references: UNIPROT:000963; UNIPARC:UP1000016B0B; GB:M92288; NID:g157019; PIDN:  
 A:Note: Sequence extracted from NCBI backbone (NCBI:108607)  
 R. Byers, T.J.; Huesin-Chahet, A.; Dubreuil, R.R.; Branton, D.; Goldstein, L.S.B.  
 J. Cell Biol. 109, 1633-1641, 1989  
 A>Title: Sequence similarity of the amino-terminal domain of *Drosophila* beta spectrin to  
 A:Reference number: A33657; MUID:90009037; PMID:2677025  
 A:Accession: A33657  
 A:Status: Preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-800 <BY2>  
 A:Cross-references: UNIPARC:UP10000173DC2; GB:M92288  
 C:Genetic8:  
 A:Gene: FlyBase:beta-Spec  
 A:Cross-references: FlyBase:FBgn0003471  
 C:Superfamily: spectrin beta-G chain; alpha-actinin actin-binding domain homology; plect  
 C:Keywords: actin binding; cytoskeleton  
 F:49-267/Domain: alpha-actinin actin-binding domain homology <ACT>  
 F:297-408/Domain: spectrin/dystrophin repeat homology <SP1>  
 F:417-522/Domain: spectrin/dystrophin repeat homology <SP2>  
 F:1698-1804/Domain: spectrin/dystrophin repeat homology <SP3>  
 F:2146-2257/Domain: plectrin repeat homology <PLK>

Query Match 5.8%; Score 316.5; DB 1; Length 2291;  
 Best Local Similarity 20.0%; Pred. No. 1e-05;  
 Matches 206; Conservative 169; Mismatches 295; Indels 359; Gaps 48;

QY 289 NASOSLWVCKEYKNTKRGVKTNTFTTSWRNGLSFCALIHFRPDLIDYKSLNPODIKEN 348  
 D 170 SAKDALLMCMKTAGYNVAVNRFTTSWRDGLAFNALIHGRPDLVQFEKLSKTNALHN 229  
 QY 349 NKAYD-GFASIGISRLLEPSDVLAIPLDKLTWTVLYQIRAHNS-----GQEL-N 398  
 D 230 LNNAFDAEDKGLAKLDAED-VFVHHPDKSITTYVYTHYFSLKQKQSTVQGRIGK 288  
 QY 399 VQJENSNSKSTYKVGNYETD-----TNSGVQDEK-----YAEISDLKR 438  
 D 289 VVGAMENDKAVHYENFTSLLKMIETTISLSGEREENSLAGVQGLAAPSRYRTIEK 348  
 QY 439 EPELQPIGAVDPLSGDDSVFVNDGVSSESSEHQTP-DDHLSPTSPYCRRTKSDTE 497  
 D 349 PPKFVE--KGNLEVL-----LFTLQSKRANNOKPYTFKGGKMSIDINKAVERLEKAEHE 401  
 QY 498 -----PKSQSSSGRTS-----GSDDPG-----S15  
 D 402 RELALRELLRQEKLEQLAARFDRKASNRKTLSENQRLVSQDFGPDLAAVEAAKKE 461  
 QY 516 -----ICSTNSTQ-----AOYLIGKRLKAKT-----LE 541  
 D 462 ALETDFAYERVAQAVACDELESERHYDKILKRDVNRMLTYILLRLARRKALE 521  
 QY 542 LS-----DLYSDRKKCMSPFICEETDEQKLTLDIGSNL-----S77  
 D 522 ISLQLOQNFQEMLYILDNMBE-----KQLMTDDYGGKLMGVEDLLQKSLVEA 571  
 QY 578 -----EKEK--LENSRSLGCRSDPE--PI-KTTSLSPTSKLGYSRDLDAKKKHA 626  
 D 572 DINILGERVAVVGNISQKF-LSDDPESTKPCDPPIYSRVQQLDAVAELRLAVEERSR 630  
 QY 627 LRQTESPDADRITLNAHSSKIVQHRLLSROE--LKERARVILLEGARDAALKAANK 684  
 D 631 L-----EESRKLMQPYWTPADENMKKEQIV-----STDEIG-- 664  
 QY 685 HNTATATPFCNRQLSDQDEERRRLRRARQLIAEASGVKMSLPSPYGAEMAKLER 744  
 D 665 HDLITTV---NLMLSKKKALE--SEITSHDPLQNVAKVG---SELITGEGFADRLIKDR 715

QY 745 SK-----ASGDENDNTEIDT-----NEEIPGFRV 769  
 D 716 LKEILAKNDHLLDLTKRRQRLKENAVYFQPLPADDDVDNMMLDLRLVSSDDV----- 769  
 QY 770 GGGDELNTLNDLTPQNSRLVDLKLKQLLEV-----QPVANSPSSAAQRAVTESSQD 825  
 D 770 -GRDE-ANVQSL-----KGGHDVDELKNVAVIDALHKQASLKLNAEKAAND----- 818  
 QY 826 MKSGTEDLKTRRLQKTRFRFPNVFVSQDSTVKTQ-LQSFQYIENRPEMKRQSIQD 884  
 D 819 -----KRLAIDNRYKELTELAK--LRQRLDALSLY-----KLMSKADGVQW 861  
 QY 885 TKKGNBEKALITENQKPSD-EVLNNGFKQTSQYVGGELAALENGQKIDTRALVAKR 943  
 D 862 IKS--KTYMLDTPMGQIDIEVEITMKRF-----GPDKENANASRAVAVNQL 908  
 QY 944 LRYLM-----DTGNTEBEEAMQEWFPMLVKKNAAL----- 975  
 D 909 ARQLLHVEHPNSDEILERQNHNLNQGWSLREKAKKDDLSAHGVQTFYIECRETISWI 968  
 QY 976 ----FRANQSLKKE-----HDLERYELNRELPMALA-----TEDMQTEAQ 1016  
 D 969 EDKGRILLETDSLMDLTVGWTLQRRLSGMDRDIAALQAKLSLREANSIEDEHPERAK 1028  
 QY 1017 KRRS-----QLLDLEVALVVKRDA-----LVRDDAQKQABEBEHLERTLEQ 1061  
 D 1029 IIRRIQIQLIWEQLTQMLKGRDSKLEAGDLHRFPLDLD-----HFQWTWLT 1077  
 QY 1062 NKGPMAKKE 1070  
 D 1078 TQTVASD 1086

RESULT 9  
 742993  
 probable spectrin beta chain - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
 C:Accession: T42993  
 R:Austin, J.; Prallie, V.; McKeown, C.  
 submitted to the EMBL Data Library, March 1998  
 A:Description: Sma-1 encodes a BH-spectrin homolog required for *C. elegans* morphogenesis  
 A:Reference number: Z22279  
 A:Accession: T42993  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-4063 <AUS>  
 A:Cross-references: UNIPROT:002425; UNIPARC:UP10000110113; EMBL:AF053496; PIDN:AAC08577.  
 C:Genetic8:  
 A:Gene: sma1  
 Query Match 5.8%; Score 316.5; DB 2; Length 4063;  
 Best Local Similarity 20.3%; Pred. No. 2.1e-05;  
 Matches 204; Conservative 150; Mismatches 353; Indels 297; Gaps 41;  
 QY 285 GRKPNASOSLWCKEYKNTKRGVKTNTFTTSWRNGLSFCALIHFRPDLIDYKSLNPOD 344  
 D 168 GERGHADADALLCORTAGYPMVRIENFTTSWRNGIAFNALIHSHRPDLVDENRLPNE 227  
 QY 345 IKENKKAYD-GFASIGISRLLEPSDVLAIPLDKLTWTVLYQIRAHNSGQD----- 396  
 D 228 HVDLNAHFVAEKGLIARLDNED-VDVTRPEKSIITVLSLYHNFPAQKTEMTGAR 286  
 QY 397 -LNVV-----QIENSNSKSTYKVGNYETDNTNSGVQDEKFFYAEISDLKRE----- 439  
 D 287 RINIVQKLMVSEFMEDDYHIASLLNMTLRVITKLTESRFPNLSNGMEENAKFNQFR 346  
 QY 440 ----PELQPIGAVDPLSGDDSVFVNDGVSSESSEHQTPDDHLSPTSPYCRRTKS 494  
 D 347 TSEKPPKYKE--KGELEAL-----FFTQTRKAKMSKQYOPQGLFMHNDIESAWAQDVA 400  
 QY 495 DTBRQ-----KSQSSSGRTSGSDDPGICSTNDS--TQAOYLIG-----S30

Db 401 ENERQVAIIAELOEKELEQLAORFHKK-----AKLRDSWLRASQVVLSEMEHGRSASQV 455  
Qy 531 KKRLKARTELESDLYVSDKKKMSPPFICEETDEOKLQTLIDISNLEKEKLEMSRLLEC 590  
Db 456 EKLTKKQQAISTDILABEDRFKMLTA--MCNELCTEKTHESDVKRGEMERLIDMTQTLT 513  
Qy 591 RSDPESPIKKTSLSPSTKLGYSYSRDLAKKKHSAIRQTESDPDARTTLNHAHSSKI 650  
Db 514 LLEQ----RRBALMSLNDL--MSLLRDIDTLNSNELYSLEPAVRNRDVCKHLIGVEDLLGK- 567  
Qy 651 VOHRLLSRQ-----BELKERARVLE-----QARRDALYAGKNTNTATPF 693  
Db 568 --HDLLDAQINAGSLSLSKLSQSANNYIRHKEBQFDVLQRLKLEDTA-----QYNTLVEL 620  
Qy 694 CNQLSDQDQEBERRQLERARQLIABARSQVMSLEPSYGEA-----AEKLEKRSKASG 749  
Db 621 C-----RSRRGLERARSLFQVQDHER-----EMWMLAEKELCTTALNSG 662  
Qy 750 D-----ENDNIEIDTNEBIPGFVVG-----GDELTNLENDL----- 782  
Db 663 DISAVPQTLLYKQVEMQTHMARSKGMLAGERLYVQNGSKEDIQRLTLQMNHRWERL 722  
Qy 783 -----DTPQONS-----KLVDL----- 794  
Db 723 RVAVDALGNWLSEARHAAQYFQDANBARSWIREKMPVLKSDDLGRDEGAASLLQRRARL 782  
Qy 795 -----KLKLLLEVOPOVANSPPSAQAKVATESSEQDKSTEDLTERLOKTERF 845  
Db 783 BEETIRAYKSDISIRLEWQSOLAN--SAFHTATTSQVQ-----TEBVNPQVEMSYNYE 835  
Qy 846 RNPVFSKDSVTKTQLOSFQYIENRPE-----MKRQR-----SIOE 883  
Db 836 GNGMRVSKGBVALLL-----KSTPEWMLAKKDGTEGYVPANYCKIVPGETVTVQ 887  
Qy 884 DTKK-----GNEEKA--ITETQRPSEDEVANKGFQTSQYVG-----ELAAL 926  
Db 888 TTQKTTTLEGNETKSSVADRHQIKISNDYRELKRLADVRRLSDNIKILRFYRCDER 947  
Qy 927 ENQKQIDTBALVEXKRLVMDTGRYTEBEEBAMQWPMVLVKNKALLIRMQOLS--LL 984  
Db 948 ERMAKEIEVSLA--DESPERHVAARFRKPKLEADMK-----TNGTQLKHINDIANDLI 1000  
Qy 985 EKEHDLERRRYELNREIRAMLAIEDMQTEAQKREQLLDELVALVNRKDALVRDLDAQ 1044  
Db 1001 SEBHGSQRKIEVQHKNINM-----WNLERLKRQGRVLE-----ATER---VADPDTT 1047  
Qy 1045 EKOAE-----EDHLEKRTLEQNKGMKKKE 1071  
Db 1048 CESAREWMLSKFEQLDRNPVDVKSQVLERDLKPLEDKIALERK 1091

RESULT 10  
T23630  
hypothetical protein R31.1 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T23630; T24242  
R:Kerhaw, J.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19774  
A:Accession: T23630  
A:status: preliminary; translated from GB/EMBL/DBJ  
A:molecule type: DNA  
A:Residues: 1-4101 <W12>  
A:Cross-references: UNIPROT:O02425; UNIPARC:UPI000017CF3B; EMBL:Z81570; PIDN:CAB04608.1;  
A:Experimental source: clone K12G11  
R:Lennard, N.  
submitted to the EMBL Data Library, July 1996  
A:Reference number: Z19862  
A:Accession: T24242  
A:status: preliminary; translated from GB/EMBL/DBJ  
A:molecule type: DNA

A:Residues: 1-4101 <W12>  
A:Cross-references: UNIPARC:UPI000017CF3B; EMBL:Z75956; PIDN:CAB00130.1; GSPDB:GN00023; (  
A:Experimental source: clone R31  
C:Genetics:  
A:Gene: CBSP.R31.1  
A:Map position: 5  
A:Insertions: 29/3; 89/3; 184/3; 312/3; 812/2; 1461/3; 1531/3; 2050/2; 2413/2; 2466/2; 3008/  
Query Match 5.8%; Score 316.5; DB 2; Length 4101;  
Best Local Similarity 20.3%; Pred. No. 2.1e-05;  
Matches 204; Conservative 150; Mismatches 353; Indels 297; Gaps 41;  
Qy 285 GRKNASOSLWCKEYTKNGYVKITNFTTSWRNGLSFSCALIHRRPDLIDYKSLNPOD 344  
Db 190 GERKHAQDALLMCGQRKTAGYPNVRIENFTTSVRNGIAFAFALLHSRPPDLVDFRNLNPN 249  
Qy 345 IKENNKAYD--GRASIGISRLPEPSMVLALIDKLTMTYLYQIARHSQGE----- 396  
Db 250 HDVNLNHAEPVAKKEIKIARLLDAED--VDVTRDEKSIITYVSLYYHHFAKQKTEMTGAR 308  
Qy 397 --LNVV-----QIENSNSKSYTKVGNVETDNNSSVDOEKFAELSDLKR----- 439  
Db 309 RIANIVKCLMVSFTMEDDYEHIASLELNMTRVITKLESRRFPNSLNGMKEEBAKNOFR 368  
Qy 440 ----DELQPIGAVDFLSQDDSVFVNDSGVSESEHQTPDDHLSPSTASPYCRTKS 494  
Db 369 TSEKPKRYK--KGELBAL-----PFIQTRKAMSKRQYPPQGLFMHDIESAMAQOLDVA 422  
Qy 495 DTEPQ-----KSQSGRTSGSDPGICSNSTD--TOAQVLLG----- 530  
Db 423 ENERQVAIIAELOEKELEQLAORFHKK-----AKLRDSWLRASQVVLSEMEHGRSASQV 477  
Qy 531 KKRLKARTELESDLYVSDKKKMSPPFICEETDEOKLQTLIDISNLEKEKLEMSRLLEC 590  
Db 478 EKLTKKQQAISTDILABEDRFKMLTA--MCNELCTEKTHESDVKRGEMERLIDMTQTLT 535  
Qy 591 RSDPESPIKKTSLSPSTKLGYSYSRDLAKKKHSAIRQTESDPDARTTLNHAHSSKI 650  
Db 536 LLEQ----RRBALMSLNDL--MSLLRDIDTLNSNELYSLEPAVRNRDVCKHLIGVEDLLGK- 589  
Qy 651 VOHRLLSRQ-----BELKERARVLE-----QARRDALYAGKNTNTATPF 693  
Db 590 --HDLLDAQINAGSLSLSKLSQSANNYIRHKEBQFDVLQRLKLEDTA-----QYNTLVEL 642  
Qy 694 CNQLSDQDQEBERRQLERARQLIABARSQVMSLEPSYGEA-----AEKLEKRSKASG 749  
Db 643 C-----RSRRGLERARSLFQVQDHER-----EMWMLAEKELCTTALNSG 684  
Qy 750 D-----ENDNIEIDTNEBIPGFVVG-----GDELTNLENDL----- 782  
Db 685 DISAVPQTLLYKQVEMQTHMARSKGMLAGERLYVQNGSKEDIQRLTLQMNHRWERL 744  
Qy 783 -----DTPQONS-----KLVDL----- 794  
Db 745 RVAVDALGNWLSEARHAAQYFQDANBARSWIREKMPVLKSDDLGRDEGAASLLQRRARL 804  
Qy 795 -----KLKLLLEVOPOVANSPPSAQAKVATESSEQDKSTEDLTERLOKTERF 845  
Db 805 BEETIRAYKSDISIRLEWQSOLAN--SAFHTATTSQVQ-----TEBVNPQVEMSYNYE 857  
Qy 846 RNPVFSKDSVTKTQLOSFQYIENRPE-----MKRQR-----SIOE 883  
Db 858 GNGMRVSKGBVALLL-----KSTPEWMLAKKDGTEGYVPANYCKIVPGETVTVQ 909  
Qy 884 DTKK-----GNEEKA--ITETQRPSEDEVANKGFQTSQYVG-----ELAAL 926  
Db 910 TTQKTTTLEGNETKSSVADRHQIKISNDYRELKRLADVRRLSDNIKILRFYRCDER 969  
Qy 927 ENQKQIDTBALVEXKRLVMDTGRYTEBEEBAMQWPMVLVKNKALLIRMQOLS--LL 984  
Db 970 ERMAKEIEVSLA--DESPERHVAARFRKPKLEADMK-----TNGTQLKHINDIANDLI 1022  
Qy 985 EKEHDLERRRYELNREIRAMLAIEDMQTEAQKREQLLDELVALVNRKDALVRDLDAQ 1044

[illegible]

R:Spelcher, D W.; Marchesi, V.T.  
 Nature 311, 177-180, 1984  
 A>Title: Erythrocyte spectrin is comprised of many homologous triple helical segments.  
 A>Reference number: A93341; MUID:84295638; PMID:6472478  
 A>Accession: B27016  
 A>Molecule type: protein  
 A>Residues: 292-324, 'X', 326-329, 'Y', 331-332, 434-533, 718-734, 'V', 736-773, 'X', 775-777, 1036-1994-1997 <SP8>  
 A>Cross-references: UNIPARC:UPI0000173DBB; UNIPARC:UPI0000173DBC; UNIPARC:UPI0000173DBD;  
 A>Note: the purified protein had a blocked amino end  
 C>Comment: Spectrin is a major structural component of the erythrocyte membrane cytoskeleton  
 C>Genetics:  
 A>Gene: GDB:SPTB  
 A>Cross-references: GDB:119602; OMIM:182870  
 A>Map position: 14q23-14q33  
 C>Superfamily: spectrin beta chain; alpha-actinin actin-binding domain homology; spectrin  
 C>Keywords: actin binding; cytoskeleton; duplication; erythrocyte; heterodimer; membrane  
 F:53-271/Domain: alpha-actinin actin-binding domain homology <ACT>  
 F:301-412/Domain: spectrin/dystrophin repeat homology <SP01>  
 F:421-526/Domain: spectrin/dystrophin repeat homology <SP02>  
 F:527-633/Domain: spectrin/dystrophin repeat homology <SP03>  
 F:636-741/Domain: spectrin/dystrophin repeat homology <SP04>  
 F:742-846/Domain: spectrin/dystrophin repeat homology <SP05>  
 F:847-953/Domain: spectrin/dystrophin repeat homology <SP06>  
 F:953-1059/Domain: spectrin/dystrophin repeat homology <SP07>  
 F:1060-1166/Domain: spectrin/dystrophin repeat homology <SP08>  
 F:1167-1272/Domain: spectrin/dystrophin repeat homology #status atypical <SP09>  
 F:1373-1377/Domain: spectrin/dystrophin repeat homology <SP10>  
 F:1378-1476/Domain: spectrin/dystrophin repeat homology <SP11>  
 F:1477-1582/Domain: spectrin/dystrophin repeat homology <SP12>  
 F:1583-1688/Domain: spectrin/dystrophin repeat homology <SP13>  
 F:1689-1795/Domain: spectrin/dystrophin repeat homology <SP14>  
 F:1796-1901/Domain: spectrin/dystrophin repeat homology <SP15>  
 F:1902-2007/Domain: spectrin/dystrophin repeat homology <SP16>  
 F:2008-2118/Domain: spectrin/dystrophin repeat homology <SP17>



```

Db      599 GYOPCDQVQIODRMSHLQCEFEELSNMAAGARTQL-EQSKLMTKFMEMDEAESEMIKEKE 657
Qy      730 LPSVGEAAAK-----LKRSKASGDENNNIETIDNEIPE--GRVVGSGDGLTYLEN 780
Db      658 -QIYSSLDYGDLTSLVILQKHKAFEDERLGLDAHLLEQIFQEAHGVNA-----RN 707
Qy      781 DLDP--EQNSKLVLDLKLKLEVPQVANSPPSSAAQKAVTSSSEODMSGTEDLATERL 838
Db      708 EFGHPQIARIKVSAAQMDQLAAFCCKNLQDAHPFQFGSADDLKMLQD----- 761
Qy      839 QKTERFRNPVPSKSTVTKTQLQSFQYIENRPEMKORSIQEDTKKNGEKAATET 898
Db      762 ---AHLRSGEDVQDQDEGATRALGKKHDFLE---ELIESRGVMEHLQ-----QA 806
Qy      899 QRKPSDEVINKGKQKQSYVVGELALENQKQIDTRALVEKRLAYLMDTGNTREEE 958
Db      807 QGPEPE-----FRD--SPDVTIRLQALRELYQGVVAQADLRQORLOBALDLYTFGETD 858
Qy      959 AMQWEMLVNKKNALIRMMQSLLEKEHDEBRREYELNRELRAMLAIED----- 1009
Db      859 A-CELM---MGEKEMLAEMEMPTLEDLVQCHRPDIQDEMKTITLQIDGVLAANSL 914
Qy      1010 -----WQ--KTBAQRRERQ----- 1026
Db      915 VESGHPRSREVKQYQDHANTRWQAFQTLVSEEREAVDLSALRVHTLCVDCERTSKMTDKT 974
Qy      1027 -----LVVALNKKDALVRD-----LDAQEKQAE--EDEHLETTLEQNK 1063
Db      975 KVESTQDPSGDLAIGTIALQKLSGLERDVAALQARVDALERESQQLMDSHPQ--KENT 1032
Qy      1064 GKMAKKEEK 1072
Db      1033 GQRQKHLE 1041

```

```

RESULT 12
T29140
hypothetical protein K11C4.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T29140
R/Pauley, A.; Gaiting, S.
submitted to the EMBL Data Library, July 1996
A/Description: The sequence of C. elegans cosmid K11C4.
A/Reference number: Z20577
A/Accession: T29140
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2326 <P>A>
A/Cross-references: UNIPARC:UPI000035886; EMBL:U64854; PIDD:AB18316.1; GSPDB:GN00023;
A/Experimental source: strain Bristol N2; clone K11C4
C/Genetics:
A/Genes: CESP:K11C4.3
A/Map position: 5
A/Intons: 105/1; 131/3; 155/3; 449/3; 2075/1; 2140/3; 2182/1; 2232/2; 2309/1
C/Superfamily: spectrin beta-G chain; alpha-actinin actin-binding domain homology; plect

```

```

Query Match 5.6%; Score 309; DB 2; Length 2326;
Best Local Similarity 20.1%; Pred. No. 2.2e-05;
Matches 201; Conservative 174; Mismatches 327; Indels 300; Gaps 44;
Qy      289 NASOSLLWCKEYKNGYKINPTTSMRNGLSFCAILHFPDLIDYKSLNPQDIKEN 348
Db      229 SAREALLLWCKQKAGPNNVNVKPFSTSWRGLAFNLIIKHHPDLVDYDNLQKSNALYN 288
Qy      349 NKRAYDGFAS-IGISRLLEPSDWLALIPDLTMTVLYOIRAHFS-----GQELNV 399
Db      289 LQSAFPTAEHQDLAKFLADE--VNVQDPBEKSIITIVVYVYHFNKLKQDNIOGRKIGK 347
Qy      400 V--QIEBNSKSYVYKGNVETDNNSSVD-----QEKYAEISDLKREPELQOPIS 447
Db      348 VINELMENDK---MINRYETLSSDLEWLNAKIQLINERHFENNLEGVQRQ-----LT 397

```

```

Qy      448 GAVDFLSQDDSVFVNDGVESESEHQTDDHLSPTASAPYC--RRTKSDTEPOKQSSG 506
Db      398 EFNVDYRQGEPRPFDER--GELEVLFTLOSARANNQRFFVREGGLIDIRANWSLE 455
Qy      507 RTSGSDDPGICSNITDSTQAVLLGKKRLKAETLE-LSDLVSDKCKKMSPPICE---- 561
Db      456 K-----AHERRELVKRELIQEKLEQLAAAF--NRKAMEMETWLTENQRL 499
Qy      562 -----ETD-----EQKLQ--TLIDGSNLEKLENSLSERSD 593
Db      500 VSQDNFNDLSVBAATKKEALETDIPAYEEVQAVAAVAGLEAENYDQAKINERKE 559
Qy      594 PESPIKKTSLPTSKLGYSYRDLAKKKAASIRQTESPPDADR--TLINAHDSKIVQ 652
Db      560 NVLQL-----WNLIFQILLARKVRLSLMAIQKIFHDMILTLDLMD-----IK 603
Qy      653 HRLISRQELKEBARVLLLEQARRDAALKA-----GNKQNTNTA-----T 691
Db      604 SRLIS--EDLGALMDVEDLLQKHALLESDDINIIGERVNNSIAQQRFRPDGPGSGYK 661
Qy      692 PFCIRQLSDQD-----ERRQLABRARQL-----IAARSGVYKSE---- 729
Db      662 PVEPQTIDERSDVLQKRYKELDLAAERKRRLBDNKLCOFMDVVALEHGIEQEQVLS 721
Qy      730 -----LPSYGEAA-----EKLKERSKASGDENNIEIDTNEIPEGFV-- 768
Db      722 STDQGRDIYVSHLAKHKAENNLDEKYLDRLDVSGELQDESIPEGSDNIPRLAET 781
Qy      769 -----VGGDELTVLENDLDPB-----QNSKLVDLK 795
Db      782 RQYINKLKELASRKERLAGVEYQFFTDADVDVRLVTLRWSESDVQDEGTVQL 841
Qy      796 LKTLLEVPQVAN-SSSSAAQKAVTSSSEODMSGTEDLATERLOCTTERFRPVPSKD 854
Db      842 LKQGDVHDELQFPDQIKVLAHAKASLPQEAR-HPDIR-QRLDTTLK----- 888
Qy      855 STVEKTOLOQSFQYIENRPEMKORSIQ-----EDTKKNGEKAATET---QR 900
Db      889 ---QKAELENTSQ-----LRKQRLIDALSILYKLYSDADSVSWIDKCKKLATLVGR 938
Qy      901 KPSEDEVINKGPDTSQYVVGELALLEN-----EQKQIDTRPAL 939
Db      939 DIREVEIMKGRFPTLBQDMNQAKVYTNVDLARQLLVNVEHPNSDDILHQNKLNARWA- 997
Qy      940 VEKRLAYLMDTGNTTEEEEMQEMFML-----VNKNALIR---RMNQLSLEKE 987
Db      998 ---QLRDMDVQKKN-ELERAHRLTEPRIDQETVTWIEBTRVLESDSALTNDLGVMK- 1052
Qy      988 HDLEREYELNRELRAMLA-----IEDMOKTBAQKRERQ--LDELVALVVKRD 1035
Db      1053 --LQRKLSMMERLUGAIOAKGLDSLHAKADDIERERQEAQAIREDIKRIHQWDIANKK- 1109
Qy      1036 ALVRDLDAQEKQAE-----EDEHLETTLEQNGKMAKKEE 1071
Db      1110 --VRBHBAKLDEAGDQLQRFRLDLDHQAULTATQROVASSEE 1149

```

```

RESULT 13
T18296
myosin heavy chain - Entamoeba histolytica
C/Species: Entamoeba histolytica
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T18296
R/Guillen, N.
submitted to the EMBL Data Library, February 1997
A/Reference number: Z18865
A/Accession: T18296
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-2139 <G>U>
A/Cross-references: UNIPROT:Q07569; UNIPARC:UPI000080203; EMBL:L03534; NID:G1850912; PIR
C/Genetics:

```



A:Gene: mhca  
 C:Superfamily: myosin heavy chain; myosin motor domain homology  
 F:91-780/Domain: myosin motor domain homology <M0>

Query Match 5.4%; Score 299; DB 2; Length 2139;  
 Best Local Similarity 20.2%; Pred. No. 5.2e-05;  
 Matches 254; Conservative 213; Mismatches 448; Indels 340; Gaps 53;

```

QY 38 FLREGKATDEDMQST-----ASLVSKQKQDIGNLDFEEDNEDDENRANQ 83
DB 659 FEEBKKDKDKEDLKKGLAEIKKGRBAENALASATKATGELAKIQDLED----- 910
QY 84 EKKAKITELINKLNF--LDEAKDLATVNSNPDPDAEL-----NPFQD----- 128
DB 911 -----KISEBSKLSAAELDKQELNLKIENL-----EEDKSELKETIDMLKQDLKQSKLG 961
QY 129 PDSEPIETETASPKRTIESFYNASNPKEVOTPOYLNPPEPAFVT-----I 177
DB 962 EDLEVEITELNSQIWLTAATVDKDKITAEWQ-----ESIDEKEDETKLGKIDILTEER 1016
QY 178 KDSPPQSTKRNIRPVMSKLYA-----DSSKTEBEELDSNPFEKSPPTPPN 227
DB 1017 KDLBQDRADVATKQDIKAKLNTKITECEDAKDIBALKEQLEDEEN-----KXK 1067
QY 228 NLVNPVQEL-----ETERRVYKAKAPAPVLSPTKTVLNTENTVSAGKDLSTSPKPSPISP 282
DB 1068 DLTEBLQOTQLKGETEKSLAAQVAATKASDERLTLSQN----- 1107
QY 283 VLGRKPNASQSLVWCKEVTYNYRGVKTNTFTTSWRNGLSPCALIHHFRPDLIDYKSLNP 342
DB 1108 -LENEKLTLYKMLTKADLEKKISGLK-----QDYEDL-- 1139
QY 343 QDINENKKAADGPAISGISRLLEPSDNLVLAIPDKLT-----VMTLYQIRAHPSGQ--- 395
DB 1140 -----EDDNKTEIGD-----LNAQRKIKELDEITKGVADVSOYLQKKEYESQIAK 1187
QY 396 ---ELNVVQIENSSEKSTYKAVNYETDT--NSSVDQEKFYAEISDKREPELQOPISGAVD 451
DB 1188 MQBEKKAIGNVKNKEKTIKEKELEIOGLQEKLBTEVYKDAE--KKGKELEKEMKAL-- 1244
QY 452 FLSDGDSVFVNDGSGVSESEHQTPDDHLSPS-----TASPYCRRTYS--DTEPQSKSQ 503
DB 1245 ---OEKKNV--ESSKNSTSEKCKKLEDNLKDTQKLDLDMTADNEKLLKAKAKDLKQALNEV 1300
QY 504 SSGRTSGSDPBGICNTDSTQAVLIGKKRLKAEITLSDLYSDKDKGMSPPICERT 563
DB 1301 QDNHEKAVADAEILNKKKAQSDKELNSLKAELEALTKAKSVESKDKOSEKKAALSBEI 1360
QY 564 DEOKLQTLDIGSNLEKEKLENSRSLSECRSDPESPICK---TSLSPSTSLGYVSRLD-- 618
DB 1361 DQANEKLNQIADLRKATADLQEANERKAEYEAQRDKLVADNKKMTTLLEIKARDEENT 1420
QY 619 -----LAKKQASLRFQTSDDPADRTTNHADHSKIVQHHLSQEBELK---ERA 666
DB 1421 YKVENYEVYLKRKEADLEBANENLDIEKKDKNKKEQYKQKLEGELKTKDLNALIAEKD 1480
QY 667 RVLEQARDAALAKRGNK---HNTNTATPFCNRQLS-----DQDERRRQLAERRA--- 714
DB 1481 SIFPAKQSDADLELNTKTEVHEHDEVAK--LNTQITLTLDNDQSAEBELNELSKDKD 1538
QY 715 ROLIAEASRGVMSLEPSYGEAAEK-----LKERSRKASGDENNDIBDNTN- 761
DB 1539 KKKISELEBOVNELESRPVGTGNADENEIKRDQADLNLKALEKGVQNNQLOQTNKEI 1598
QY 762 -----EIPGFFVGGDELITNLENDLITPEQNS-----KLVLDKL 796
DB 1599 KAKNDLTSKIEITE-----NEKKLENAKKQLEQDKDADKAVSQTIKRGLEBEV 1651
QY 797 KKL-LFVQP---QVANSFSSAAQKAVTSSSQDKMSGTFEDLRTERL-----OKTTER 844
DB 1652 KKLTLTEIQALKFQI--NAPSSVAOEBEKQRLSDIALKLEQLEOEFTTAAANAEARKKIQ 1710
QY 845 FRNPFVPS--KDSVTKTQLQSFSGYIENRPKMKQORSIQEDTKKGNNEKAITETQRP 903

```

```

DB 1711 ELDEVKFVLBDVTNQREKL--VAKNSEDAID--SLKEBKKALEDEIKETIDDNKLS 1765
QY 904 E-----DEVIAKGFQDTSQYVVGELA---ALENEQK----- 931
DB 1766 EIDSLDKRYNALLDSKQSDVSMKEKFD-----ELKVTQDLLETKKNHAETMRKLG 1818
QY 932 QIDTRALVETKRLRYL-----MDTGRTSEEBAMQEWFWLVNKKGALIRRN--QL 981
DB 1819 RLEKKAQVQRLNLELQKNDLDAQEKAKATKDRADDELKSLNLELDVYQDLDRKQD 1878
QY 982 SLSEKHD--LSRRYELN-----ELRAMLAIEDWQTEAQKRE--QLLLDE 1026
DB 1879 DLADKEDLATLQDKYKTLVQKSVPSRIGEMQEQDLERAKGRAKQKQKQAVEKKIQE 1938
QY 1027 LVALVN-----KRDALVRDLDAQEKAEEDHLEHTLEFQNGKRAKK--EKK 1073
DB 1939 LQENDNPFEEYKETADKRINTLSAQKDLQKELEK--ERGLQDSEKEVQRLRVK 1992

```

## RESULT 14

528916  
 dystrophin - mouse  
 N:Alternate names: duchenne muscular dystrophy protein  
 C:Species: Mus musculus (house mouse)  
 C>Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text\_change 09-Jul-2004  
 C:Accession: S28916; S10922; C43837; B40134  
 R:Bies, R.D.; Phelps, S.F.; Cotez, W.D.; Roberts, R.; Caskey, C.T.; Chamberlain, J.S.  
 Nucleic Acids Res. 20, 1725-1731, 1992  
 A>Title: Human and murine dystrophin mRNA transcripts are differentially expressed during  
 A:Reference number: S28916; MUID:92253376; PMID:1579466  
 A:Accession: S28916  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-3678 <BIE>  
 A:Cross-references: UNIPROT:P11531; UNIPARC:UPI0000279E7; EMBL:M68859  
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1991  
 R:Koenig, M.; Hoffman, R.P.; Bertelson, C.J.; Monaco, A.P.; Feener, C.; Kunkel, L.M.  
 Cell 50, 509-517, 1987  
 A>Title: Complete cloning of the Duchenne muscular dystrophy (DMD) cDNA and preliminary  
 A:Reference number: A90897; MUID:8727512; PMID:3607877  
 A:Accession: B27162  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-201 <KOB>  
 A:Cross-references: UNIPARC:UPI0000177686  
 R:Medel, U.; Zuk, D.; Elnat, P.; Zeelon, E.; Levy, Z.; Neuman, S.; Yaffe, D.  
 Nature 337, 76-78, 1989  
 A>Title: Duchenne muscular dystrophy gene product is not identical in muscle and brain.  
 A:Reference number: S06461; MUID:89082658; PMID:2909892  
 A:Accession: S10922  
 A:Status: translation not shown  
 A:Molecule type: mRNA  
 A:Residues: 1-106 <NUD>  
 A:Cross-references: UNIPARC:UPI0000177687; EMBL:X14183  
 R:Rapaport, D.; Lederfein, D.; den Dunnen, J.T.; Grootscholten, P.M.; Van Ommen, G.J.; R.  
 Differentiation 49, 187-193, 1992  
 A>Title: Characterization and cell type distribution of a novel, major transcript of the  
 A:Reference number: A43837; MUID:92316332; PMID:1377655  
 A:Accession: C43837  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 3069-3181 <RAP>  
 A:Cross-references: UNIPARC:UPI0000177688  
 A:Note: sequence extracted from NCBI backbone  
 R:Hoffman, B.P.; Monaco, A.P.; Feener, C.C.; Kunkel, L.M.  
 Science 238, 347-350, 1987  
 A>Title: Conservation of the Duchenne muscular dystrophy gene in mice and humans.  
 A:Reference number: A40134; MUID:88018015; PMID:3659917  
 A:Accession: B40134  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 300-676, 'F', 678-1390 <HOF>



Db 1446 ARDAVEHLTKSLA-----DVESQVSVQNO-----EKDALGRL 1478  
QY 533 RLKAEITELSDLYVSDKKOMSPFICETDEQLQTLIDIGSNLE-----KE 580  
Db 1479 ALLQOE-----RDK-----LIVEMDKSILENOSLGSCSLKALGLTDPKE 1521  
QY 581 KL-ENSRSLCSDPDESPICKTSLSPTSK-----LGYSYRDLDLAKK----- 622  
Db 1522 KLMELESVRCSKIAESTEMQEKHELOKEYEVLLQSTENNSNEARIQHVYESVROEKQ 1581  
QY 623 -KHASLRQTESDPPDARTTLNHADHSKIYQHRLLSROEBLKERAR-----VLLBOA 673  
Db 1582 EYVATLRSASPSDKRERKQLOAE-----QEMEMEKERKRFKSKQKILELE 1630  
QY 674 RDDALKA-----GNKHTNTATPFQNRQLSDQDE---ERRQLRE-----RAQLIAE 720  
Db 1631 BENDRLRAEQPVGGANESMEALLSSNASLKEBELERTLEYKTLSEKEFEALMAERNTLSE 1690  
QY 721 ARSGVQNSELPSYGEAAEKLKERESKASGDENDNIETNEBIEPQVYVGGDELNTLEN 780  
Db 1691 ETRNLKL-----QVEAQELQASLETTEKSDEPK-DVIEVYEA-VVGKSQE----- 1735  
QY 781 DLDTPQNSKLVDLKLLKLEVPQVANSPPSAAQAVTE--SSQDMKSGTEDL----- 833  
Db 1736 -QDSISENAKLEDAEATLL-----ANS---AKPGVSETPSSHDDINNYLQQLDQLKG 1783  
QY 834 -----RTERLOKTERFRNPV---VPSKOSTVR---KTQLOSPSOYIENRPEM 875  
Db 1784 RIABELEMEKQORELSOTLENEKNALLTOISAKOSELKLEBEVAKINMLNQIOE--EL 1841  
QY 876 KRORSIOEDTKKGNBE-----KAAITETQRPKSDEVLNKGFKDTSOVVGEI 923  
Db 1842 SRVTKLKTAEERKODLEERLNMQLAELNGSIGNTYQVYTDQIKNQOLESBMQNLKRCV 1901  
QY 924 AALENEQKQIDTRALVERKLR--YL-----MDTGRNTEBEBAMQEWPMLVNKKNAL-- 974  
Db 1902 SELEBEKQOLVBKTKVSEBIRKEYMEKIQGAQKGPSSKIHAKELQE--LLEKQOEYKQ 1959  
QY 975 -----IRRMQLSLEKHEHDLERYEELNRELRAML-----AIEDMQTEAQKRE 1020  
Db 1960 LQKDCIRYLRISALEK--TVKALIEFVHTESQKDLATKGNLAQAVEHHKKAQAEISSP 2016  
QY 1021 QLLADE-----LVALLVNRDALVPRDLDAQEQAEEDHLEERTLEFONKGMAKKEEK 1072  
Db 2017 KILHDDTQSEARVLAADNLKKELOSNKESIKSQIKQKDDLLRLRQAEK-HRREKK 2075

Search completed: May 20, 2006, 17:15:34  
Job time : 59 secs

**THIS PAGE BLANK (USPTO)**

This Page is inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record

## BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLORED OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REPERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images  
problems checked, please do not report the  
problems to the IFW Image Problem Mailbox**

**THIS PAGE IS BLANK**